

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 19:47:31 ; Search time 144.742 Seconds

(without alignments)  
371.762 Million cell updates/sec

Title: US-10-054-873-7

Perfect score: 797

Sequence: 1 MFPTIPLSRLFDNMLSLAHR.....IVEGCTSTCSLYOLENYCN 150

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729239 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Geneseq\_23Sep04:\*  
2: geneseqp1980s:\*  
3: geneseqp1930s:\*  
4: geneseqp2000s:\*  
5: geneseqp2001s:\*  
6: geneseqp2002s:\*  
7: geneseqp2003as:\*  
8: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	797	100.0	150	2	AAy42861
2	555.5	69.7	107	2	AAy42861
3	470	59.0	92	2	AAy42861
4	470	59.0	92	2	AAy42861
5	470	59.0	92	2	AAy42861
6	466	58.5	188	8	ADi47330
7	466	58.5	188	8	ADi47330
8	466	58.5	188	8	ADi47330
9	466	58.5	188	8	ADi47330
10	466	58.5	188	8	ADi47330
11	466	58.5	188	8	ADi47330
12	466	58.5	188	8	ADi47330
13	466	58.5	188	8	ADi47330
14	466	58.5	188	8	ADi47330
15	466	58.5	188	8	ADi47330
16	466	58.5	188	8	ADi47330
17	466	58.5	188	8	ADi47330
18	466	58.5	188	8	ADi47330
19	466	58.5	188	8	ADi47330
20	466	58.5	188	8	ADi47330
21	466	58.5	188	8	ADi47330
22	466	58.5	188	8	ADi47330
23	466	58.5	188	8	ADi47330
24	466	58.5	188	8	ADi47330
25	466	58.5	188	8	ADi47330

26	464	58.2	191	8	ADL89541	ADL89541 Human mod
27	463	58.1	191	5	ABG94860	ABG94860 Human gro
28	463	58.1	191	5	ABG94977	ABG94977 Human gro
29	463	58.1	191	7	ADK41914	ADK41914 Human gro
30	463	58.1	191	8	ADL89532	ADL89532 Human gro
31	463	58.1	191	8	ADL89532	ADL89532 Human mod
32	463	58.1	191	8	ADL89532	ADL89532 Human mod
33	463	58.1	191	8	ADL89532	ADL89532 Human mod
34	463	58.1	191	8	ADL89532	ADL89532 Human mod
35	462	58.0	144	2	AAy04396	AAy04396 Human alb
36	462	58.0	191	5	ABG94975	ABG94975 Human gro
37	462	58.0	191	5	ABG94975	ABG94975 Human gro
38	462	58.0	191	7	ADK41912	ADK41912 Human gro
39	462	58.0	191	7	ADK41913	ADK41913 Human gro
40	462	58.0	191	8	ADL89547	ADL89547 Human mod
41	462	58.0	191	8	ADL89547	ADL89547 Human mod
42	462	58.0	191	8	ADL89547	ADL89547 Human mod
43	462	58.0	191	8	ADL89547	ADL89547 Human mod
44	461	57.8	191	2	AAy04396	AAy04396 Human bet
45	461	57.8	191	2	AAy04396	AAy04396 Natural h

#### ALIGNMENTS

RESULT 1  
AAy42861  
ID AAy42861 standard; protein; 150 AA.  
XX  
AC AAy42861;  
XX  
DT 19-JUN-2000 (first entry)  
XX  
DE Chimeric protein, SEQ ID 7.  
XX  
KW Insulin; precursor; growth hormone; chaperone; intramolecular; folding;  
KW conformation; chimeric protein; cleavable; recombinant; production;  
KW yield.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
PN WO990302-A1.  
XX  
PD 07-OCT-1999.  
XX  
PF 31-MAR-1998; 98WO-CN000052.  
XX  
PR 31-MAR-1998; 98WO-CN000052.  
XX  
PA (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.  
XX  
PI Gan Z;  
XX  
PT WPI; 1999-610839/52.  
XX  
PS Claim 14; Page 30-31; 46pp; English.  
XX  
This sequence represents a chimeric protein, which contains an N-terminal fragment of human growth hormone (hGH) of the sequence given in AAy42856, a cleavable peptide linker (AAy42857), and a human insulin precursor comprising insulin A and B chains (AAy42859). The hGH portion of the chimeric protein acts as an intramolecular chaperone (IMC) for the insulin precursor, enabling it to fold correctly. The cleavable peptide linker has a C-terminal Arg residue which enables the hGH portion of the chimeric protein to be removed after folding has taken place. Production of recombinant human insulin via an hGH-proinsulin chimeric protein can provide human insulin with correctly linked cysteine bridges with fewer necessary procedural steps, and hence resulting in a higher yield of human insulin. The IMC sequences not only protect insulin sequences from

CC intracellular degradation by a microorganism host, but also promote the  
 CC folding of the fused insulin precursor, facilitate the solubility of the  
 CC fusion protein and decrease the intermolecular interactions among the  
 CC fusion proteins, thus allowing folding of the fused insulin precursor at  
 CC commercially useful high concentrations. The procedural steps of cyanogen  
 CC bromide cleavage, oxidative sulphydrololysis and related purification steps  
 CC can thus be eliminated, along with the use of high concentrations of  
 CC mercapтан or the use of hydrophobic absorbent resins

XX Sequence 150 AA;

Query Match 100.0%; Score 797; DB 2; Length 150;

Best Local Similarity 100.0%; Pred. No. 1.4e-45;  
 Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQEFEEAVYPRCKXYSFLONPOTSISPSISIP 60

DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQEFEEAVYPRCKXYSFLONPOTSISPSISIP 60

DB 61 TPSNREETQCKSNLELRISILLIQSWLEPVLGTGPRFVNOHLCGSHLVEALYVCGER 120

DB 61 TPSNREETQCKSNLELRISILLIQSWLEPVLGTGPRFVNOHLCGSHLVEALYVCGER 120

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

CC hGH-proinsulin chimeric protein can provide human insulin with correctly  
 CC linked cysteine bridges with fewer necessary procedural steps, and hence  
 CC resulting in a higher yield of human insulin. The IMC sequences not only  
 CC protect insulin sequences from intracellular degradation by a  
 CC microorganism host, but also promote the folding of the fused insulin  
 CC precursor, facilitate the solubility of the fusion protein and decrease  
 CC the intermolecular interactions among the fusion proteins, thus allowing  
 CC folding of the fused insulin precursor at commercially useful high  
 CC concentrations. The procedural steps of cyanogen bromide cleavage,  
 CC oxidative sulphydrololysis and related purification steps can thus be  
 CC eliminated, along with the use of high concentrations of mercapтан or the  
 CC use of hydrophobic absorbent resins

XX Sequence 107 AA;

Query Match 69.7%; Score 555.5; DB 2; Length 107;

Best Local Similarity 71.3%; Pred. No. 1.1e-29;  
 Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

XX

DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQEFEEAVYPRCKXYSFLONPOTSISPSISIP 60

DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYQEFEEAVYPRCKXYSFLONPOTSISPSISIP 60

DB 61 TPSNREETQCKSNLELRISILLIQSWLEPVLGTGPRFVNOHLCGSHLVEALYVCGER 120

DB 61 TPSNREETQCKSNLELRISILLIQSWLEPVLGTGPRFVNOHLCGSHLVEALYVCGER 120

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

DB 121 GFYTPKTRGIVEQCCTSIQSLYOLENYCN 150

RESULT 3  
 AAY42856  
 ID AAY42856 standard; protein; 92 AA.

XX AAY42856;

XX 19-JAN-2000 (first entry)

XX hGH-mini-proinsulin chimeric protein.

XX Insulin; precursor; growth hormone; chaperone; intramolecular; folding;

XX conformation; chimeric protein; cleavable; recombinant; production;

XX yield.

XX Homo sapiens.

XX WO9950302-A1.

XX 07-OCT-1999.

XX 31-MAR-1998; 98WO-CN000052.

XX 31-MAR-1998; 98WO-CN000052.

XX (TONG-) TONGHUA GANTECH BIOTECHNOLOGY LTD.

XX Gan Z;

XX WPI; 1999-610839/52.

XX New chimeric proteins containing human growth hormone fragment, used

XX particularly for the production of human insulin.

XX Claim 5; Page 28; 46pp; English.

XX

XX This sequence represents an N-terminal fragment of human growth hormone

XX (hGH) which is a component of a chimeric protein (AAY42861) which also

XX contains a human insulin precursor (AAY42859). The hGH portion of the

XX chimeric protein acts as an intramolecular chaperone (IMC) for the

XX insulin precursor, enabling it to fold correctly. A cleavable peptide

XX linker with a C-terminal Arg residue (AAY42857) enables the hGH portion

XX

CC of the chimeric protein to be removed after folding has taken place.  
 CC production of recombinant human insulin via an hgh-proinsulin chimeric  
 CC protein can provide human insulin with correctly linked cysteine bridges  
 CC with fewer necessary procedural steps, and hence resulting in a higher  
 CC yield of human insulin. The IYK sequences not only protect insulin  
 CC sequences from intracellular degradation by a microorganism host, but  
 CC also promote the folding of the fused insulin precursor, facilitate the  
 CC solubility of the fusion protein and decrease the intermolecular  
 CC interactions among the fusion proteins, thus allowing folding of the  
 CC fused insulin precursor at commercially useful high concentrations. The  
 CC procedural steps of cyanogen bromide cleavage, oxidative sulphydrolysis  
 CC and related purification steps can thus be eliminated, along with the use  
 CC of high concentrations of mercapran or the use of hydrophobic absorbent  
 CC resins

SQ Sequence 92 AA;

Query Match 59.0%; Score 470; DB 2; Length 92;  
 Best Local Similarity 100.0%; Pred. NO. 4.3e-24;  
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSLRPFDMNMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPOTSLSFSSESIP 60  
 Db 1 MFPTPLSLRPFDMNMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPOTSLSFSSESIP 60  
 QY 61 TPSNREETOQKSNLELRISLLILIOSWLEPVQ 92  
 Db 61 TPSNREETOQKSNLELRISLLILIOSWLEPVQ 92

RESULT 4  
 AAM92265  
 ID AAM92265 standard; protein; 134 AA.

AC AAM92265;  
 DT 08-JUN-1999 (first entry)

DE Human anti-angiogenic peptide 16K hgh Met-1Prol33.

XX Human; anti-angiogenic; prolactin; placental lactogen; hpl; angiogenesis;  
 KM growth hormone; hgh; hgh-V; capillary endothelial cell proliferation;  
 KM placental vasculatization; pregnancy; treatment; angiogenic disease;  
 KM tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;  
 KM arthritis; atherosclerotic plaques; corneal graft neovascularisation;  
 KM wound healing; proliferative retinopathy; macular degeneration; trachoma;  
 KM granuloma; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;  
 KM psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;  
 KM ulcer; leukaemia; reproductive disorder; contraceptive agent;  
 KM gene therapy; pre-eclampsia; intrauterine growth retardation;  
 KM placental dysfunction.

OS Homo sapiens.

XX W09851323-A1.

XX 19-NOV-1998.

XX 12-MAY-1998; 98WC-US009691.

XX 13-MAY-1997; 97US-0046394P.

XX (RBGC) UNIV CALIFORNIA.

XX Weiner RI, Martial JA, Struman I, Taylor R;

XX WPI; 1999-045192/04.

XX N-PSDB; AAX01707.

PT New anti-angiogenic peptides - comprise N-terminal fragments of human  
 PT placental lactogen, human growth hormone, growth hormone variant or human  
 PT prolactin.

PS Claim 4; Page 49-50; 87pp; English.

XX This invention describes novel human anti-angiogenic peptides derived  
 CC from 10 to 150 consecutive amino acids selected from the N-terminal end  
 CC of human placental lactogen (hpl), human growth hormone (hgh), growth  
 CC hormone variant (hgh-V), or human prolactin. Such peptides (i) inhibit  
 CC capillary endothelial cell proliferation and organisation (ii) inhibit  
 CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at  
 CC least one specific receptor which does not bind an intact full length  
 CC hgh, hpl, prolactin or hgh-V. The invention also describes a method for  
 CC diagnosing a probable abnormality of placental vascularisation during  
 CC pregnancy. The peptides can be used for treating an angiogenic disease in  
 CC a subject, for inhibiting tumour formation or growth in a patient or for  
 CC modulating vascularisation of a patient's placenta. In particular, the  
 CC peptides can be used for preventing or treating e.g. malignant tumours,  
 CC angiofibroma, arteriovenous malformation, arthritic such as rheumatoid  
 CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,  
 CC delayed wound healing, proliferative retinopathy such as diabetic  
 CC retinopathy, macular degeneration, granulations such as those occurring  
 CC in haemophilic joints, inappropriate vascularisation in wound healing  
 CC such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular  
 CC tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis,  
 CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,  
 CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,  
 CC and choriocarcinoma. They can also be used as contraceptive agents. DNA  
 CC encoding the peptides can be used in gene therapy. The measurement of  
 CC abnormal levels of N-terminal fragments of hgh, hgh-V, prolactin or hpl  
 CC can be used in assays for impairment of vascular development associated  
 CC with pre-eclampsia, intrauterine growth retardation, and placental  
 CC dysfunction

SQ Sequence 134 AA;

Query Match 59.0%; Score 470; DB 2; Length 134;  
 Best Local Similarity 100.0%; Pred. NO. 5.8e-24;  
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSLRPFDMNMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPOTSLSFSSESIP 60  
 Db 1 MFPTPLSLRPFDMNMLRAHRLHQLAFDTYQFEFEAYIPKQKYSFLQNPOTSLSFSSESIP 60  
 QY 61 TPSNREETOQKSNLELRISLLILIOSWLEPVQ 92  
 Db 61 TPSNREETOQKSNLELRISLLILIOSWLEPVQ 92

RESULT 5  
 ABG94861  
 ID ABG94861 standard; protein; 191 AA.

AC ABG94861;

DT 03-DEC-2002 (first entry)

DE Human growth hormone mutant hPRL (111-129).

XX Growth hormone; placental lactogen; prolactin; active domain; hgh;  
 KM structure-function relationship; segment-substituted polypeptide; mutant;  
 KM mutin.

XX Homo sapiens.

XX Synthetic.

XX US6428954-B1.

XX 06-AUG-2002.

XX 06-JUN-1995; 95US-00483039.

XX 28-OCT-1998; 88US-00264611.  
 XX 26-OCT-1998; 89US-00428066.  
 XX 27-APR-1992; 92US-00875204.

PR 13-OCT-1992; 92US-00960227.  
 PR 02-FEB-1994; 94US-00190723.

XX (GENE) GENENTECH INC.

XX PI Wells JA, Cunningham BC;

XX WPI; 2002-696875/75.

XX Identifying active domains within cloned polypeptides of known amino acid  
 PT sequence by substituting analog segments into the parent polypeptide is  
 PT useful to determine the relationship between structure and function.

PS Example 1; Page: 86pp; English.

CC The invention relates to identifying an unknown active domain in a region  
 CC of known amino acid sequence in a parent polypeptide e.g. human growth  
 CC hormone (hGH) which has been cloned and has a pre-identified biological  
 CC activity, where the active domain interacts with a target when the parent  
 CC polypeptide is in its native-folded form and the interaction is  
 CC responsible for the biological activity comprising: (a) comparing the  
 CC amino acid sequence or polypeptide structure in the region of known amino  
 CC acid sequence of hGH with the amino acid sequence of an analogue  
 CC structure in a region of known amino acid sequence of an analogue  
 CC polypeptide (e.g. prolactin, placental lactogen or porcine growth  
 CC hormone) which has at least 15% homology with hGH alpha-carbon  
 CC coordinates within about 2-3.5 angstroms of hGH alpha-carbon coordinates  
 CC for about 60% of the analogue sequence, where any interaction of the  
 CC analogue with the target is different from target interaction with hGH;  
 CC (b) substituting DNA encoding an analogous polypeptide segment from the  
 CC analogue into DNA encoding the full length hGH, and expressing a segment-  
 CC substituted polypeptide; (c) contacting the segment-substituted  
 CC polypeptide with the target to determine interaction; (d) repeating steps  
 CC (b) and (c) with a second analogous polypeptide segment; and (e)  
 CC comparing the difference between activity of the first and second segment  
 CC -substituted polypeptides as an indication of the location of the unknown  
 CC active domain in hGH. The method is useful for determining the  
 CC relationship between structure and function of known polypeptide  
 CC sequences. The present sequence is that of human growth hormone mutant  
 CC substituted with residues from an hGH analogue (prolactin, placental  
 CC lactogen or porcine growth hormone). Note: The present sequence is not  
 CC shown in the specification but was created by the indexer using the  
 CC mature hGH sequence and information contained in the specification

XX Sequence 191 AA;

SO Query Match 59.0%; Score 470; DB 5; Length 191;  
 Best Local Similarity 69.2%; Pred. No. 7.8e-24;  
 Matches 101; Conservative 8; Mismatches 19; Indels 18; Gaps 3;

QY 2 PPTPLSRPLPDNMLRAHRLHQLAFPTYQFFEEAYIPKQKXSFLLNPQTSLSFSSSIP 61  
 DB 1 PPTPLSRPLPDNMLRAHRLHQLAFPTYQFFEEAYIPKQKXSFLLNPQTSLSFSSSIP 60

QY 62 PSNNEETQOKSNLELRLISLLIQSWLEPVLGTGPRFVNOHLCGS-----H 108  
 DB 61 PSNNEETQOKSNLELRLISLLIQSWLEPVLGTGPRFVNOHLCGS-----H 119

QY 109 LVEALYVCGERGFFYPKTRGIVEQ 134  
 DB 120 LVEALYVCGERGFFYPKTRGIVEQ 141

RESULT 6  
 ADI47330  
 ID ADI47330 standard; protein; 188 AA.

XX ADI47330;

XX 22-APR-2004 (first entry)

XX Plasmid pOAI1A1 amino acid sequence SEQ ID NO:18.

KM Multimer assembly; DNA sequence; amplification cassette;  
 KM monomer sequence; restriction pair member; diagnostic protein;  
 KM therapeutic protein.

XX Synthetic.

XX WO2004007687-A2.

XX 22-JAN-2004.

XX 16-JUL-2003; 2003WO-US022216.

XX 16-JUL-2002; 2002US-0396466P.

XX (BUSSE/) BUSSELL S.

XX Busse1 S;

XX WPI; 2004-122926/12.

XX N-PSDB; ADI47329.

XX Multimer assembly of DNA sequences comprising an amplification cassette  
 PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
 PT terminus and 3' RPM at its 3' terminus.

PS Example 2; SEQ ID NO 18; 163pp; English.

XX The present invention describes a multimer assembly of DNA sequences (1)  
 CC comprising at least one amplification cassette (AC) having at least one  
 CC monomer sequence whose polymerization is desired, and a 5' restriction  
 CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
 CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
 CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
 CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
 CC RPM site of AC. (1) can be used for expressing a diagnostic protein or  
 CC therapeutic protein. In (1), the diagnostic protein and therapeutic  
 CC protein is a cytokine, a growth factor, a transcription factor, a receptor,  
 CC ligand, an enzyme, an inhibitor, a transcription factor, a transcription  
 CC factor, a DNA replication protein, an activator, a chaperonin, or an  
 CC antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,  
 CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
 CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
 CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
 CC granulocyte-macrophage colony-stimulating factor, leukemia inhibitory  
 CC factor, tumour necrosis factor, lymphotxin, platelet-derived growth  
 CC factor, fibroblast growth factor, vascular endothelial cell growth  
 CC factor, epidermal growth factor, transforming growth factor-beta,  
 CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
 CC oncostatin M, amphiregulin, müllerian-inhibiting substance, B-cell growth  
 CC factor, macrophage migration inhibiting factor, endostatin, or  
 CC angiotensin. The present sequence is used in the exemplification of the  
 CC present invention.

XX Sequence 188 AA;

SO Query Match 58.5%; Score 466; DB 8; Length 188;  
 Best Local Similarity 70.5%; Pred. No. 1.4e-23;  
 Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MPTPLSRPLPDNMLRAHRLHQLAFPTYQFFEEAYIPKQKXSFLLNPQTSLSFSSSIP 60  
 DB 1 MPTPLSRPLPDNMLRAHRLHQLAFPTYQFFEEAYIPKQKXSFLLNPQTSLSFSSSIP 60

QY 61 TPSNNEETQOKSNLELRLISLLIQSWLEPVLGTGPRFVNOHLCGS-----HLV 110  
 DB 61 TPSNNEETQOKSNLELRLISLLIQSWLEPVLGTGPRFVNOHLCGS-----HLV 119

QY 111 EALYVCG-ERGFFYPKTRGIVEQ 134  
 DB 120 EGIOTLWLRLEDG---SPRTGQIFKQ 142

RESULT 7



AAP90129  
 ID AAP90129 standard; protein; 192 AA.  
 AC AAP90129;  
 XX  
 DT 24-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 06-FEB-1996 (revised)  
 DT 01-NOV-1989 (first entry)  
 XX  
 DE Human growth hormone.  
 OS Homo sapiens; (Human).  
 XX  
 PN JP01144981-A.  
 PD 07-JUN-1989.  
 XX  
 PF 02-DEC-1987; 87JP-00304937.  
 XX  
 PR 02-DEC-1987; 87JP-00304937.  
 XX  
 PA (WAKT ) WAKUNAGA SEIYAKU KK.  
 DR WPI; 1989-209284/29.  
 DR N-PSDB; AAN90269.  
 XX  
 PT Recombinant vector cong. fused protein aminoacid coding - composed of  
 PT growth hormone or its polypeptide deriv. and foreign protein.  
 XX  
 PS Disclosure; Fig 1; 19pp; Japanese.  
 XX  
 CC The invention consists of a vector cong. a fusion protein which is  
 CC formed by ligating, downstream of a promoter, hGH or a deriv. (pref.  
 CC formed by substen. of Met-14 with Ieu) and a foreign protein. Stability  
 CC of the vector in the host is greatly increased so the protein yield is  
 CC higher. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 24-OCT-  
 CC 2003 to standardise OS field)  
 XX  
 SQ Sequence 192 AA;  
 XX  
 QY Query Match 58.5%; Score 466; DB 1; Length 192;  
 Db Best Local Similarity 70.5%; Pred. No.1.4e-23;  
 Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;  
 QY 1 MFPTIPLSRLFDNMLRAHRLHQLAFPTYOEFEBAYIPKQKYSFLONPOTSLSFSSSIP 60  
 Db 1 MFPTIPLSRLFDNMLRAHRLHQLAFPTYOEFEBAYIPKQKYSFLONPOTSLSFSSSIP 60  
 QY 61 TPNSREETQOKSNIELIRISILLIQSWLEPVQGTGRFVNQHLCS-----HLV 110  
 Db 61 TPNSREETQOKSNIELIRISILLIQSWLEPVQGTGRFVNQHLCS-----HLV 110  
 QY 111 EALYVCG--ERGFYPTKTRGIVEQ 134  
 Db 120 EGIOTLMGRLEDG---SPRTGQIFKQ 142  
 XX  
 RESULT 8  
 AAW92264  
 ID AAW92264 standard; protein; 192 AA.  
 AC AAW92264;  
 XX  
 DT 08-JUN-1999 (first entry)  
 XX  
 DE Human anti-angiogenic peptide hGH Met-1Phe191.  
 XX  
 DE Human; anti-angiogenic; prolactin; Placental lactogen; hPL; angiogenesis;  
 KM growth hormone; hGH; hGH-V; capillary endothelial cell proliferation;  
 KM placental vascularisation; pregnancy; treatment; angiogenic disease;

KM tumour; inhibitor; malignant; angiofibroma; arteriovenous malformation;  
 KM arthritis; atherosclerotic plaques; corneal graft neovascularisation;  
 KM wound healing; proliferative retinopathy; macular degeneration; trachoma;  
 KM granulation; glaucoma; ocular; uveitis; fracture; Osler-Weber syndrome;  
 KM psoriasis; fibroplasia; scleroderma; Kaposi's sarcoma; vascular adhesion;  
 KM ulcer; leukaemia; reproductive disorder; contraceptive agent;  
 KM gene therapy; pre-eclampsia; intrauterine growth retardation;  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09851323-A1.  
 PD 19-NOV-1998.  
 XX  
 PF 12-MAY-1998; 98MO-US009691.  
 XX  
 PR 13-MAY-1997; 97US-0046394P.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PT Weiner RI, Martial JA, Struman I, Taylor R;  
 XX  
 DR WPI; 1999-045192/04.  
 DR N-PSDB; AAX01706.  
 XX  
 PT New anti-angiogenic peptides - comprise N-terminal fragments of human  
 PT placental lactogen, human growth hormone, growth hormone variant or human  
 PT prolactin.  
 XX  
 PS Example 3; Page 49; 87pp; English.  
 XX  
 CC This invention describes novel human anti-angiogenic peptides derived  
 CC from 10 to 150 consecutive amino acids selected from the N-terminal end  
 CC of human placental lactogen (hPL), human growth hormone (hGH), growth  
 CC hormone variant (hGH-V), or human prolactin. Such peptides (i) inhibit  
 CC capillary endothelial cell proliferation and organisation (ii) inhibit  
 CC angiogenesis in chick chorioallantoic membrane and (iii) binds to at  
 CC least one specific receptor which does not bind an intact full length  
 CC hGH, hPL, prolactin or hGH-V. The invention also describes a method for  
 CC diagnosing a probable abnormality of placental vascularisation during  
 CC pregnancy. The peptides can be used for treating an angiogenic disease in  
 CC a subject, for inhibiting tumour formation or growth in a patient or for  
 CC modulating vascularisation of a patient's placenta. In particular, the  
 CC peptides can be used for preventing or treating e.g. malignant tumours,  
 CC angiofibroma, arteriovenous malformation, arthritic such as rheumatoid  
 CC arthritis, atherosclerotic plaques, corneal graft neovascularisation,  
 CC delayed wound healing, proliferative retinopathy such as diabetic  
 CC retinopathy, macular degeneration, granulations such as those occurring  
 CC in haemophilic joints, inappropriate vascularisation in wound healing  
 CC such as hypertrophic scars or keloid scars, neovascular glaucoma, ocular  
 CC tumour, uveitis, non-union fractures, Osler-Weber syndrome, psoriasis,  
 CC pyogenic glaucoma, retrolental fibroplasia, scleroderma, solid tumours,  
 CC Kaposi's sarcoma, trachoma, vascular adhesions, chronic varicose ulcers,  
 CC leukaemia, and reproductive disorders such as follicular and luteal cysts  
 CC and choroidcarcinoma. They can also be used as contraceptive agents. DNA  
 CC encoding the peptides can be used in gene therapy. The measurement of  
 CC abnormal levels of N-terminal fragments of hGH, hGH-V, prolactin or hPL  
 CC can be used in assays for impairment of vascular development associated  
 CC with pre-eclampsia, intrauterine growth retardation, and placental  
 CC dysfunction  
 XX  
 SQ Sequence 192 AA;  
 XX  
 QY Query Match 58.5%; Score 466; DB 2; Length 192;  
 Db Best Local Similarity 70.5%; Pred. No.1.4e-23;  
 Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;  
 QY 1 MFPTIPLSRLFDNMLRAHRLHQLAFPTYOEFEBAYIPKQKYSFLONPOTSLSFSSSIP 60  
 Db 1 MFPTIPLSRLFDNMLRAHRLHQLAFPTYOEFEBAYIPKQKYSFLONPOTSLSFSSSIP 60  
 QY 61 TPNSREETQOKSNIELIRISILLIQSWLEPVQGTGRFVNQHLCS-----HLV 110

```

DB 61 TPSSNEETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNYDLKCLE 119
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134
DB 120 EGIQTLMGRLSDG---SPRTGQIFKQ 142

RESULT 9
AD147320
ID AD147320 standard; protein; 192 AA.
AC AD147320;
XX 22-APR-2004 (first entry)
XX Plasmid p0A0 amino acid sequence SEQ ID NO:8.
XX multimer assembly; DNA sequence; amplification cassette;
XX monomer sequence; restriction pair member; diagnostic protein;
XX therapeutic protein.
XX Synthetic.
XX WO2004007687-A2.
XX 22-JAN-2004.
XX 16-JUL-2003; 2003WO-US022216.
XX 16-JUL-2002; 2002US-0396466P.
XX (BUSS/) BUSSELL S.
XX Buswell S;
XX WPI; 2004-122926/12.
XX N-PSDB; AD147319.
XX Multimer assembly of DNA sequences comprising an amplification cassette
XX having monomer sequences and 5' restriction pair member (RPM) at its 5'
XX terminus and 3' RPM at its 3' terminus.
XX Example 1; SEQ ID NO 8; 163pp; English.
XX The present invention describes a multimer assembly of DNA sequences (I)
XX comprising at least one amplification cassette (AC) having at least one
XX monomer sequence whose polymerisation is desired, and a 5' restriction
XX pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and
XX one or more of following: (a) 3'-terminal cassette comprising 3' specific
XX sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal
XX cassette comprising 5' specific sequence and 3' RPM site fused to a 5'
XX RPM site of AC. (I) can be used for expressing a diagnostic protein or
XX therapeutic protein. In (I), the diagnostic protein and therapeutic
XX protein is a cytokine, a growth factor, a hormone, a receptor, a receptor
XX ligand, an enzyme, an inhibitor, a transcription factor, a translation
XX factor, a DNA replication factor, an activator, a chaperonin, or an
XX antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,
XX IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,
XX IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,
XX colony-stimulating factor-1, granulocyte colony-stimulating factor,
XX granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory
XX factor, tumour necrosis factor, lymphotoxin, platelet-derived growth
XX factor, fibroblast growth factor, vascular endothelial cell growth
XX factor, epidermal growth factor, transforming growth factor-beta,
XX transforming growth factor-alpha, thrombopoietin, stem cell factor,
XX oncostatin M, amphiregulin, melanin-inhibiting substance, B-cell growth
XX factor, macrophage migration inhibiting factor, endostatin, or
XX angiotensin. The present sequence is used in the exemplification of the
XX present invention.
XX Sequence 192 AA;

```

```

Query Match 58.5%; Score 466; DB 8; Length 192;
Best Local Similarity 70.5%; Pred. No. 1,4e-23;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MPEPTPLSLRFDNAMRAHRLHQLADPTQSEBEAYIPBEQKXSLQNPQTSLSSEIP 60
DB 1 MPEPTPLSLRFDNAMRAHRLHQLADPTQSEBEAYIPBEQKXSLQNPQTSLSSEIP 60
QY 61 TPSSNEETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNYDLKCLE 119
DB 61 TPSSNEETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNYDLKCLE 119
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134
DB 120 EGIQTLMGRLSDG---SPRTGQIFKQ 142

RESULT 10
AD147390
ID AD147390 standard; protein; 192 AA.
AC AD147390;
XX 22-APR-2004 (first entry)
XX Plasmid p0A5A amino acid sequence SEQ ID NO:78.
XX multimer assembly; DNA sequence; amplification cassette;
XX monomer sequence; restriction pair member; diagnostic protein;
XX therapeutic protein.
XX Synthetic.
XX WO2004007687-A2.
XX 22-JAN-2004.
XX 16-JUL-2003; 2003WO-US022216.
XX 16-JUL-2002; 2002US-0396466P.
XX (BUSS/) BUSSELL S.
XX Buswell S;
XX WPI; 2004-122926/12.
XX N-PSDB; AD147389.
XX Multimer assembly of DNA sequences comprising an amplification cassette
XX having monomer sequences and 5' restriction pair member (RPM) at its 5'
XX terminus and 3' RPM at its 3' terminus.
XX Example 12; SEQ ID NO 78; 163pp; English.
XX The present invention describes a multimer assembly of DNA sequences (I)
XX comprising at least one amplification cassette (AC) having at least one
XX monomer sequence whose polymerisation is desired, and a 5' restriction
XX pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and
XX one or more of following: (a) 3'-terminal cassette comprising 3' specific
XX sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal
XX cassette comprising 5' specific sequence and 3' RPM site fused to a 5'
XX RPM site of AC. (I) can be used for expressing a diagnostic protein or
XX therapeutic protein. In (I), the diagnostic protein and therapeutic
XX protein is a cytokine, a growth factor, a hormone, a receptor, a receptor
XX ligand, an enzyme, an inhibitor, a transcription factor, a translation
XX factor, a DNA replication factor, an activator, a chaperonin, or an
XX antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,
XX IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,
XX IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,
XX colony-stimulating factor-1, granulocyte colony-stimulating factor,
XX granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory
XX factor, tumour necrosis factor, lymphotoxin, platelet-derived growth
XX factor, fibroblast growth factors, vascular endothelial cell growth

```

CC factor, epidermal growth factor, transforming growth factor-beta,  
 CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
 CC oncostatin M, amphiregulin, mulierian-inhibiting substance, B-cell growth  
 CC factor, macrophage migration inhibiting factor, endostatin, or  
 CC angiotatin. The present sequence is used in the exemplification of the  
 CC present invention.

XX Sequence 192 AA;

SO Query Match 58.5%; Score 466; DB 8; Length 192;  
 Best Local Similarity 70.5%; Pred. No. 1.4e-23;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MPTTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLNPTQTSLSFSRSP 60  
 Db 1 MPTTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLNPTQTSLSFSRSP 60

QY 61 TFSNNEETQOKSNLELRISLLIQSWLEPVQLGTGPRFVNHLCGS-----HLV 110  
 Db 61 TFSNNEETQOKSNLELRISLLIQSWLEPVQLGTGPRFVNHLCGS-----HLV 110

QY 111 EALYIVCG--ERGFYTPKTRGIVEQ 134  
 Db 120 EGIQTLMGRLDGG--SPRTGQIRPKQ 142

RESULT 11

AD147398  
 ID AD147398 standard; protein; 192 AA.

XX AD147398;

DT 22-APR-2004 (first entry)

DE Nmer amplification cassette amino acid sequence SEQ ID NO:86.

XX multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

XX therapeutic protein.

OS Synthetic.

PN WO2004007687-A2.

PD 22-JAN-2004.

PF 16-JUL-2003; 2003WO-US022216.

PR 16-JUL-2002; 2002US-0396466P.

PA (BUSELL) BUSELL S.

PI Bussell S;

DR WPI: 2004-122926/12.

DR P-PsDB; AD147397.

PT Multimer assembly of DNA sequences comprising an amplification cassette

PT having monomer sequences and 5' restriction pair member (RPM) at its 5'

PT terminus and 3' RPM at its 3' terminus.

XX Claim 115; SEQ ID NO 86; 163pp; English.

CC The present invention describes a multimer assembly of DNA sequences (I)  
 CC comprising at least one amplification cassette (AC) having at least one  
 CC monomer sequence whose polymerization is desired, and a 5' restriction  
 CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
 CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
 CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
 CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
 CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
 CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
 CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor

CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
 CC factor, a DNA replication factor, an activator, a chaperonin, or an  
 CC antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,  
 CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
 CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
 CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
 CC granulocyte-macrophage colony-stimulating factor, leukemia inhibitory  
 CC factor, tumour necrosis factor, lymphotoxin, platelet-derived growth  
 CC factor, fibroblast growth factor, vascular endothelial cell growth  
 CC factor, epidermal growth factor, transforming growth factor-beta,  
 CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
 CC oncostatin M, amphiregulin, mulierian-inhibiting substance, B-cell growth  
 CC factor, macrophage migration inhibiting factor, endostatin, or  
 CC angiotatin. The present sequence is used in the exemplification of the  
 CC present invention.

XX Sequence 192 AA;

SO Query Match 58.5%; Score 466; DB 8; Length 192;  
 Best Local Similarity 70.5%; Pred. No. 1.4e-23;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MPTTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLNPTQTSLSFSRSP 60  
 Db 1 MPTTPLSRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLNPTQTSLSFSRSP 60

QY 61 TFSNNEETQOKSNLELRISLLIQSWLEPVQLGTGPRFVNHLCGS-----HLV 110  
 Db 61 TFSNNEETQOKSNLELRISLLIQSWLEPVQLGTGPRFVNHLCGS-----HLV 110

QY 111 EALYIVCG--ERGFYTPKTRGIVEQ 134  
 Db 120 EGIQTLMGRLDGG--SPRTGQIRPKQ 142

RESULT 12

AD147354  
 ID AD147354 standard; protein; 193 AA.

XX AD147354;

DT 22-APR-2004 (first entry)

DE Plasmid pOR31A amino acid sequence SEQ ID NO:42.

XX multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

XX therapeutic protein.

OS Synthetic.

PN WO2004007687-A2.

PD 22-JAN-2004.

PF 16-JUL-2003; 2003WO-US022216.

PR 16-JUL-2002; 2002US-0396466P.

PA (BUSELL) BUSELL S.

PI Bussell S;

DR WPI: 2004-122926/12.

DR N-PsDB; AD147353.

PT Multimer assembly of DNA sequences comprising an amplification cassette

PT having monomer sequences and 5' restriction pair member (RPM) at its 5'

PT terminus and 3' RPM at its 3' terminus.

XX Example 7; SEQ ID NO 42; 163pp; English.

CC The present invention describes a multimer assembly of DNA sequences (I)

comprising at least one amplification cassette (AC) having at least one monomer sequence whose polymerisation is desired, and a 5' restriction pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and one or more of following: (a) 3'-terminal cassette comprising 3' specific sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal cassette comprising 5' specific sequence and 3' RPM site fused to a 5' RPM site of AC. (1) can be used for expressing a diagnostic protein or therapeutic protein. In (1), the diagnostic protein and therapeutic protein is a cytokine, a growth factor, a hormone, a receptor, a receptor ligand, an enzyme, an inhibitor, a transcription factor, a translation factor, a DNA replication factor, an activator, a chaperonin, or an antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta, IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin, colony-stimulating factor-1, granulocyte colony-stimulating factor, granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory factor, tumour necrosis factor, lymphotoxin, platelet-derived growth factor, fibroblast growth factors, vascular endothelial cell growth factor, epidermal growth factor, transforming growth factor-beta, transforming growth factor-alpha, thrombopoietin, stem cell factor, oncostatin M, amphiregulin, muellerian-inhibiting substance, B-cell growth factor, macrophage migration inhibiting factor, endostatin, or angiotensin. The present sequence is used in the exemplification of the present invention.

Sequence 193 AA;

Query Match 58.5%; Score 466; DB 8; Length 193;  
Best Local Similarity 70.5%; Pred. No. 1,4e-23;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYOEFEENAVYIKKQKXSFLOQPSLFSSESIP 60  
DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYOEFEENAVYIKKQKXSFLOQPSLFSSESIP 60  
QY 61 TPNREBETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDL 119  
DB 61 TPNREBETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDL 119  
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 120 EGIQTLMGRLDQ---SPRTGQIFKQ 142

RESULT 13

ID ADI47384 standard; protein; 206 AA.

AC ADI47384;

DT 22-APR-2004 (first entry)

DE Plasmid pOA43A insert amino acid sequence SEQ ID NO:72.

XX multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

XX therapeutic protein.

OS Synthetic.

PN WO2004007687-A2.

PD 22-JAN-2004.

PF 16-JUL-2003; 2003WO-US022216.

PR 16-JUL-2002; 2002US-0396466P.

PA (BUSS/) BUSS/LL S.

PI Bussell S;

XX Bussell S;

XX Bussell S;

XX Bussell S;

DR WPI; 2004-122926/12.

DR P-PSDB; ADI47383.

XX Multimer assembly of DNA sequences comprising an amplification cassette  
PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
PT terminus and 3' RPM at its 3' terminus.

XX Example 11; SEQ ID NO 72; 163bp; English.

XX The present invention describes a multimer assembly of DNA sequences (1) comprising at least one amplification cassette (AC) having at least one monomer sequence whose polymerisation is desired, and a 5' restriction pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and one or more of following: (a) 3'-terminal cassette comprising 3' specific sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal cassette comprising 5' specific sequence and 3' RPM site fused to a 5' RPM site of AC. (1) can be used for expressing a diagnostic protein or therapeutic protein. In (1), the diagnostic protein and therapeutic protein is a cytokine, a growth factor, a hormone, a receptor, a receptor ligand, an enzyme, an inhibitor, a transcription factor, a translation factor, a DNA replication factor, an activator, a chaperonin, or an antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta, IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin, colony-stimulating factor-1, granulocyte colony-stimulating factor, granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory factor, tumour necrosis factor, lymphotoxin, platelet-derived growth factor, fibroblast growth factors, vascular endothelial cell growth factor, epidermal growth factor, transforming growth factor-beta, transforming growth factor-alpha, thrombopoietin, stem cell factor, oncostatin M, amphiregulin, muellerian-inhibiting substance, B-cell growth factor, macrophage migration inhibiting factor, endostatin, or angiotensin. The present sequence is used in the exemplification of the present invention.

Sequence 206 AA;

Query Match 58.5%; Score 466; DB 8; Length 206;  
Best Local Similarity 70.5%; Pred. No. 1.5e-23;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYOEFEENAVYIKKQKXSFLOQPSLFSSESIP 60  
DB 1 MFPTPLSRFLPDNMLRAHRLHQLAFDTYOEFEENAVYIKKQKXSFLOQPSLFSSESIP 60  
QY 61 TPNREBETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDL 119  
DB 61 TPNREBETQOKSNLELRISLLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDL 119  
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 120 EGIQTLMGRLDQ---SPRTGQIFKQ 142

RESULT 14

ID ADI47363 standard; protein; 391 AA.

AC ADI47363;

DT 22-APR-2004 (first entry)

DE Plasmid pOA31F2 insert amino acid sequence SEQ ID NO:51.

XX multimer assembly; DNA sequence; amplification cassette;

KW monomer sequence; restriction pair member; diagnostic protein;

XX therapeutic protein.

OS Synthetic.

PN WO2004007687-A2.

PD 22-JAN-2004.

```

PF      16-JUL-2003; 2003MO-US022216.
XX
XX      16-JUL-2002; 20020US-0396466P.
XX
XX      (BUSEL/) BUSELL S.
XX
XX      Busell S;
XX
XX      WPI; 2004-122926/12.
XX      N-PSDB; ADI47362.
XX
XX      Multimer assembly of DNA sequences comprising an amplification cassette
PT      having monomer sequences and 5' restriction pair member (RPM) at its 5'
PT      terminus and 3' RPM at its 3' terminus.
XX
XX      Example 8; SEQ ID NO 51; 163bp; English.
XX
XX      The present invention describes a multimer assembly of DNA sequences (I)
CC      comprising at least one amplification cassette (AC) having at least one
CC      monomer sequence whose polymerisation is desired, and a 5' restriction
CC      pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and
CC      one or more of following: (a) 3'-terminal cassette comprising 3' specific
CC      sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal
CC      cassette comprising 5' specific sequence and 3' RPM site fused to a 5'
CC      RPM site of AC. (1) can be used for expressing a diagnostic protein or
CC      therapeutic protein. In (II), the diagnostic protein and therapeutic
CC      protein is a cytokine, a growth factor, a hormone, a receptor, a receptor
CC      ligand, an enzyme, an inhibitor, a transcription factor, a translation
CC      factor, a RNA replication factor, an activator, a chaperonin, or an
CC      antibody. The therapeutic protein is interferon (IFN) alpha, IFN-beta,
CC      IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,
CC      IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,
CC      colony-stimulating factor-1, granulocyte colony-stimulating factor,
CC      granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory
CC      factor, tumour necrosis factor, lymphotoxin, platelet-derived growth
CC      factor, fibroblast growth factors, vascular endothelial cell growth
CC      factor, epidermal growth factor, transforming growth factor-beta,
CC      transforming growth factor-alpha, thrombopoietin, stem cell factor,
CC      oncostatin M, amphiregulin, mullerian-inhibiting substance, B-cell growth
CC      factor, macrophage migration inhibiting factor, endostatin, or
CC      angiotstatin. The present sequence is used in the exemplification of the
CC      present invention.
XX
XX
SQ      Sequence 391 AA;
XX
XX
XX      Query Match          58.5%; Score 466; DB 8; Length 391;
XX      Best Local Similarity 70.5%; Pred.No. 2.6e-23;
XX      Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4
XX
XX      1 MPEPTILPSLRFNDAMRAHRLHQLAFDPVYQCFEENAYIPKQKYSFLQNPOTSLSEFESIP 60
XX      1 MPEPTILPSLRFNDAMRAHRLHQLAFDPVYQCFEENAYIPKQKYSFLQNPOTSLSEFESIP 60
XX      1 MPEPTILPSLRFNDAMRAHRLHQLAFDPVYQCFEENAYIPKQKYSFLQNPOTSLSEFESIP 60
XX      1 MPEPTILPSLRFNDAMRAHRLHQLAFDPVYQCFEENAYIPKQKYSFLQNPOTSLSEFESIP 60
XX
XX      61 TPNREBETQOKSNLDELIRISLLISLMLEPQVLGGRPFVNHQCGS-----HLV 110
XX      61 TPNREBETQOKSNLDELIRISLLISLMLEPQVLGGRPFVNHQCGS-----HLV 110
XX      61 TPNREBETQOKSNLDELIRISLLISLMLEPQVLGGRPFVNHQCGS-----HLV 110
XX      61 TPNREBETQOKSNLDELIRISLLISLMLEPQVLGGRPFVNHQCGS-----HLV 110
XX
XX      111 EALYLWCG--ERGFYTPKTRGIVQ 134
XX      111 EALYLWCG--ERGFYTPKTRGIVQ 134
XX      120 EGIQTLMGRLDQ--SPRTGQIFQ 142
XX      120 EGIQTLMGRLDQ--SPRTGQIFQ 142
XX
XX      RESULT 15
XX      ADI47344
XX      ID ADI47344 standard; protein; 574 AA.
XX
XX      ADI47344;
XX
XX      22-APR-2004 (first entry).
XX
XX      Plasmid pDA1D insert amino acid sequence SEQ ID NO:32.
XX
XX      multimer assembly; DNA sequence; amplification cassette;
XX

```

monomer sequence; restriction pair member; diagnostic protein;  
 therapeutic protein.  
 XX Synthesis.  
 OS Synthetic.  
 KW WO2004007687-A2.

XX  
 XX  
 XX PD 22-JAN-2004.

XX PF 16-JUL-2003; 2003WO-US0222216.  
 XX PR 16-JUL-2002; 2002US-0396466P.  
 XX PA (BUSSELL) BUSELL S.  
 XX PI Bussell S;  
 XX DR WIPI: 2004-122926/12.  
 XX DR N-PSTDB; ADI47343.

XX Multimer assembly of DNA sequences comprising an amplification cassette  
 PT having monomer sequences and 5' restriction pair member (RPM) at its 5'  
 PT terminus and 3' RPM at its 3' terminus.

PS Claim 67; SEQ ID NO 32; 163pp; English.

XX The present invention describes a multimer assembly of DNA sequences (I)  
 CC comprising at least one amplification cassette (AC) having at least one  
 CC monomer sequence whose polymerization is desired, and a 5' restriction  
 CC pair member (RPM) at its 5' terminus and 3' RPM at its 3' terminus, and  
 CC one or more of following: (a) 3'-terminal cassette comprising 3' specific  
 CC sequence and 5' RPM site fused to a 3' RPM site of AC; or (b) 5'-terminal  
 CC cassette comprising 5' specific sequence and 3' RPM site fused to a 5'  
 CC RPM site of AC. (I) can be used for expressing a diagnostic protein or  
 CC therapeutic protein. In (I), the diagnostic protein and therapeutic  
 CC protein is a cytokine, a growth factor, a hormone, a receptor, a receptor  
 CC ligand, an enzyme, an inhibitor, a transcription factor, a translation  
 CC factor, a DNA replication factor, an activator, a chaperonin, or an  
 CC antibody. The therapeutic protein is interferon (IFN alpha, IFN-beta,  
 CC IFN-gamma, interleukin (IL)-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8,  
 CC IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, erythropoietin,  
 CC colony-stimulating factor-1, granulocyte colony-stimulating factor,  
 CC granulocyte-macrophage colony-stimulating factor, leukaemia inhibitory  
 CC factor, tumor necrosis factor, lymphotoxin, platelet-derived cell growth  
 CC factor, fibroblast growth factors, vascular endothelial cell growth  
 CC factor, epidermal growth factor, transforming growth factor-beta,  
 CC transforming growth factor-alpha, thrombopoietin, stem cell factor,  
 CC oncostatin M, amphiregulin, multiplex-inhibiting substance, B-cell growth  
 CC factor, macrophage migration inhibiting factor, endostatin, or  
 CC angiostatin. The present sequence is used in the exemplification of the  
 CC present invention.

XX Sequence 574 AA:

SQ

Query Match	58.5%;	Score 466;	DB 8;	Length 574;
Best Local Similarity	70.5%;	Pred. No. 3.5e-22;		
Matches 103;	Conservative 7;	Mismatches 20;	Indels 16;	Gaps 4

OY 1 MFPTITSLRTPNNAMARHRLHQALAFDYQGFEEAVYIPKROKYSFLONPOTSLSFSSIP 60  
 DB 1 MFPTITSLRTPNNAMARHRLHQALAFDYQGFEEAVYIPKROKYSFLONPOTSLCFSSISIP 60

OY 61 TSPNRRETOOKSNLELIRISLLILSQWLEPVQLGTGPPFNQHTCGS-----HLV 110  
 DB 61 TSPNRRETOOKSNLELIRISLLILSQWLEPVQF-LRSVFANSLVGYASDSGNVYDLKLDE 119

OY 111 EALYLVCG--ERGFYETPKRTGIIEQ 134  
 DB 120 EGIGTLMGRLEDG---SPRTQGIFRQ 142

Search completed: November 2, 2004, 20:11:49  
 Job time : 145.742 secs

Sat Nov 6 18:59:31 2004

us-10-054-873-7.rag

Sat Nov 6 18:59:32 2004

us-10-054-873-7.rai

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 2, 2004, 20:02:41, Search time 36.5314 Seconds  
(without alignments)  
272.306 Million cell updates/sec

Title: US-10-054-873-7  
Perfect score: 797  
Sequence: 1 MPTPLSLRLFDNMLRAHR.....IVECCSISLXQLENYCN 150

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfilest.pep.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	466	58.5	192	1 US-08-093-383-1	Sequence 1, Appl1
2	461	57.8	191	3 US-09-284-878-5	Sequence 5, Appl1
3	461	57.8	191	4 US-09-462-941-1	Sequence 1, Appl1
4	461	57.8	194	2 US-08-383-621-4	Sequence 4, Appl1
5	461	57.8	194	3 US-08-459-906-4	Sequence 4, Appl1
6	461	57.8	217	3 US-08-589-028-10	Sequence 10, Appl1
7	461	57.8	217	3 US-08-784-582-10	Sequence 10, Appl1
8	461	57.8	217	3 US-08-785-271-10	Sequence 10, Appl1
9	461	57.8	217	3 US-08-785-628-11	Sequence 11, Appl1
10	461	57.8	217	3 US-09-284-878-1	Sequence 1, Appl1
11	461	57.8	217	4 US-09-929-918-9	Sequence 9, Appl1
12	461	57.8	241	4 US-09-424-6208-25	Sequence 25, Appl1
13	461	57.8	245	4 US-09-280-030-66	Sequence 66, Appl1
14	461	57.8	274	4 US-08-784-582-71	Sequence 71, Appl1
15	461	57.8	360	4 US-08-784-582-73	Sequence 73, Appl1
16	460	57.7	191	4 US-09-554-451-1	Sequence 1, Appl1
17	455	57.1	191	3 US-09-462-941-1	Sequence 1, Appl1
18	455	57.1	191	4 US-09-554-451-3	Sequence 3, Appl1
19	455	57.1	217	1 US-08-187-756C-4	Sequence 4, Appl1
20	455	57.1	217	1 US-08-469-466-51	Sequence 51, Appl1
21	455	57.1	217	1 US-08-469-658-51	Sequence 51, Appl1
22	455	57.1	217	2 US-08-710-324A-4	Sequence 4, Appl1
23	455	57.1	217	4 US-09-411-657-4	Sequence 4, Appl1
24	454	57.0	400	4 US-09-420-819-37	Sequence 37, Appl1
25	454	57.0	401	4 US-09-420-819-37	Sequence 36, Appl1
26	448	56.2	191	3 US-08-800-215C-18	Sequence 18, Appl1
27	446	56.0	191	3 US-08-800-215C-16	Sequence 16, Appl1

28	446	56.0	191	3 US-08-800-215C-20	Sequence 20, Appl1
29	365.5	45.9	176	3 US-08-791-728-1	Sequence 1, Appl1
30	365.5	45.9	176	3 US-08-990-774-1	Sequence 1, Appl1
31	359.5	45.1	176	3 US-08-791-728-2	Sequence 2, Appl1
32	359.5	45.1	176	3 US-08-990-774-2	Sequence 2, Appl1
33	343	43.0	168	6 5424199-3	Patent No. 5424199
34	334.5	42.0	198	1 US-08-187-756C-5	Sequence 5, Appl1
35	334.5	42.0	198	2 US-08-710-324A-5	Sequence 5, Appl1
36	334.5	42.0	198	4 US-09-411-657-5	Sequence 5, Appl1
37	316.5	39.7	96	1 US-08-160-376A-5	Sequence 5, Appl1
38	316.5	39.7	96	1 US-08-389-487-8	Sequence 8, Appl1
39	314	39.4	137	1 US-08-400-256-39	Sequence 39, Appl1
40	314	39.4	137	3 US-08-975-365-38	Sequence 39, Appl1
41	309.5	38.8	146	1 US-08-400-256-48	Sequence 48, Appl1
42	309.5	38.8	146	3 US-08-975-365-48	Sequence 48, Appl1
43	309	38.8	145	1 US-08-400-256-45	Sequence 45, Appl1
44	309	38.8	145	3 US-08-975-365-45	Sequence 45, Appl1
45	306.5	38.5	191	1 US-08-468-824-8	Sequence 8, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-093-383-1  
Sequence 1, Application US/08093383  
Patent No. 5489529  
GENERAL INFORMATION:  
APPLICANT: DeBoer, Herman A.  
APPLICANT: Heynker, Herbert L.  
APPLICANT: Seeburg, Peter H.  
TITLE OF INVENTION: DNA for Expression of Bovine Growth Hormone  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/093,383  
FILING DATE: 14-JUL-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/615827  
FILING DATE: 28-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/198624  
FILING DATE: 05-APR-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/632361  
FILING DATE: 19-JUL-1984  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/303687  
FILING DATE: 18-SEP-1981  
ATTORNEY/AGENT INFORMATION:  
NAME: Johnston, Sean A.  
REGISTRATION NUMBER: 935,910  
REFERENCE/DOCKET NUMBER: 46C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3562  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 amino acids  
TYPE: amino acid



TOPOLOGY: linear  
US-08-093-383-1

Query Match  
Best Local Similarity 58.5%; Score 466; DB 1; Length 192;  
Best Local Similarity 70.5%; Pred. No. 1.8e-42;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPOTSLSFSSISPT 60  
DB 1 MPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPOTSLSFSSISPT 60  
QY 61 TPNNEETQOKSNLELRLISILLIQSWLEPVQGLGTFPYNCHLCS-----HLV 110  
DB 61 TPNNEETQOKSNLELRLISILLIQSWLEPVQGLGTFPYNCHLCS-----HLV 110  
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 120 GIQTLMGRLDGG--SPRTGQIFKQ 142

#### RESULT 2

US-09-284-878-5  
Sequence 5, Application US/09284878  
Patent No. 6342375

#### GENERAL INFORMATION:

APPLICANT: Olazaran, Martha Guerrero  
APPLICANT: Saldaña, Hugo Barrera  
APPLICANT: Saldaña, Jose Maria Valdez  
TITLE OF INVENTION: Genetically Modified Methylophilic P. pastoris Yeast for the  
FILE REFERENCE: 1829, 0010000  
CURRENT APPLICATION NUMBER: US/09/284, 878  
CURRENT FILING DATE: 1999-07-21  
PRIOR APPLICATION NUMBER: PCT/NX97/00033  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patent Ver. 2.1  
SEQ ID NO 5  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-284-878-5

Query Match  
Best Local Similarity 57.8%; Score 461; DB 3; Length 191;

Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPOTSLSFSSISPT 61  
DB 1 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPOTSLSFSSISPT 60  
QY 62 TPNNEETQOKSNLELRLISILLIQSWLEPVQGLGTFPYNCHLCS-----HLV 111  
DB 61 TPNNEETQOKSNLELRLISILLIQSWLEPVQGLGTFPYNCHLCS-----HLV 111  
QY 112 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 120 GIQTLMGRLDGG--SPRTGQIFKQ 141

#### RESULT 3

US-09-462-941-1  
Sequence 1, Application US/09462941  
Patent No. 6608183

#### GENERAL INFORMATION:

APPLICANT: Cox III, George N  
APPLICANT: Bolder Biotechnology, Inc.  
TITLE OF INVENTION: Derivatives of Growth Hormone and Related Proteins  
FILE REFERENCE: 4152-1-PUS  
CURRENT APPLICATION NUMBER: US/09/462, 941  
CURRENT FILING DATE: 2000-01-14  
PRIOR APPLICATION NUMBER: 60/052, 516  
PRIOR FILING DATE: 1997-07-14

NUMBER OF SEQ ID NOS: 41  
SOFTWARE: Patent Ver. 2.0  
SEQ ID NO 1  
LENGTH: 191  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-462-941-1

Query Match  
Best Local Similarity 57.8%; Score 461; DB 4; Length 191;  
Best Local Similarity 70.3%; Pred. No. 6.1e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPOTSLSFSSISPT 61  
DB 1 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPOTSLSFSSISPT 60  
QY 62 TPNNEETQOKSNLELRLISILLIQSWLEPVQGLGTFPYNCHLCS-----HLV 111  
DB 61 TPNNEETQOKSNLELRLISILLIQSWLEPVQGLGTFPYNCHLCS-----HLV 111  
QY 112 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 120 GIQTLMGRLDGG--SPRTGQIFKQ 141

#### RESULT 4

US-08-383-621-4  
Sequence 4, Application US/08383621  
Patent No. 5951972

#### GENERAL INFORMATION:

APPLICANT: Daley, Michael J.  
APPLICANT: Buckwalter, Brian L.  
APPLICANT: Cady, Susan M.  
APPLICANT: Shieh, Hong-Ming  
APPLICANT: Bohlen, Peter  
APPLICANT: Seddon, Andrew P.  
TITLE OF INVENTION: Stabilization Of Somatotropins And Other  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. Estelle J. Tsevdos  
STREET: 1937 West Main Street, P.O. Box 60  
CITY: Stamford  
STATE: Connecticut  
COUNTRY: U.S.A.  
ZIP: 06904-0060

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/383, 621  
FILING DATE: 06-FEB-1995  
CLASSIFICATION: 514

#### PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/766, 142  
FILING DATE: 25-SEP-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Tsevdos, Estelle J.  
REGISTRATION NUMBER: 31,145  
REFERENCE/DOCKET NUMBER: 31,278-01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 203-321-2756  
TELEFAX: 203-321-2971  
TELEX: 203-710-474-4059

#### INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 194 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: Protein  
US-08-383-621-4

Query Match 57.8%; Score 461; DB 2; Length 194;  
Best Local Similarity 70.3%; Pred. No. 6.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTPLSRPLFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLQNPOTSLSFSISPT 61  
DB 4 FFTPLSRPLFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLQNPOTSLSFSISPT 63

QY 62 PSNRRETOQKSNLELIRISILLIQSWLEPVQGTGFRFVNOHLCGS-----HLYE 111  
DB 64 PSNRRETOQKSNLELIRISILLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDLLE 122

QY 112 ALYVCG--ERGFYTPKTRGIVEQ 134  
DB 123 GIQTLMGRLBDG---SPRTGQIFKQ 144

RESULT 5  
US-08-459-906-4  
Sequence 4, Application US/08459906  
Patent No. 6010999  
GENERAL INFORMATION:  
APPLICANT: Daley, Michael J.  
APPLICANT: Buckwalter, Brian L.  
APPLICANT: Casey, Susan M.  
APPLICANT: Shieh, Hong-Ming  
APPLICANT: Bohlen, Peter  
APPLICANT: Seddon, Andrew P.  
TITLE OF INVENTION: Stabilization of Somatotropins and Other  
TITLE OF INVENTION: Proteins by Modification of Cysteine Residues  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: U.S.A.  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,906  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Webster, Darryl L.  
REGISTRATION NUMBER: 34,276  
REFERENCE/DOCKET NUMBER: 31,278-03  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-831-3247  
TELEFAX: 201-831-3305  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 194 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: protein  
US-08-459-906-4

Query Match 57.8%; Score 461; DB 3; Length 194;  
Best Local Similarity 70.3%; Pred. No. 6.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTPLSRPLFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLQNPOTSLSFSISPT 61  
DB 4 FFTPLSRPLFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLQNPOTSLSFSISPT 63

QY 62 PSNRRETOQKSNLELIRISILLIQSWLEPVQGTGFRFVNOHLCGS-----HLYE 111  
DB 64 PSNRRETOQKSNLELIRISILLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDLLE 122

DB 64 PSNRRETOQKSNLELIRISILLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDLLE 122

QY 112 ALYVCG--ERGFYTPKTRGIVEQ 134  
DB 123 GIQTLMGRLBDG---SPRTGQIFKQ 144

RESULT 6  
US-08-589-028-10  
Sequence 10, Application US/08589028  
Patent No. 6087129  
GENERAL INFORMATION:  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Halban, Philippe  
APPLICANT: No. 6087129mington, Karl D.  
APPLICANT: Clark, Samuel A.  
APPLICANT: Thigpen, Anice E.  
APPLICANT: Quade, Christian  
APPLICANT: Kruse, Fred  
TITLE OF INVENTION: Recombinant Expression of Proteins From  
TITLE OF INVENTION: Secretary Cell Lines  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P. O. Box 4433  
CITY: Houston  
STATE: TX  
COUNTRY: USA  
ZIP: 77210-4433  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/589,028  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 47,642  
REFERENCE/DOCKET NUMBER: UTSD:426\HYL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-589-028-10

Query Match 57.8%; Score 461; DB 3; Length 217;  
Best Local Similarity 70.3%; Pred. No. 7.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTPLSRPLFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLQNPOTSLSFSISPT 61  
DB 27 FFTPLSRPLFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLQNPOTSLSFSISPT 86

QY 62 PSNRRETOQKSNLELIRISILLIQSWLEPVQGTGFRFVNOHLCGS-----HLYE 111  
DB 87 PSNRRETOQKSNLELIRISILLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDLLE 145

QY 112 ALYVCG--ERGFYTPKTRGIVEQ 134  
DB 146 GIQTLMGRLBDG---SPRTGQIFKQ 167

RESULT 7  
US-08-784-582-10  
Sequence 10, Application US/08784582

Patent No. 6110707

## GENERAL INFORMATION:

APPLICANT: Newgard, Christopher B.  
APPLICANT: Halban, Philippe A.  
APPLICANT: No. 610707/Minington, Karl D.  
APPLICANT: Clark, Samuel A.  
APPLICANT: Thigpen, Anice E.  
APPLICANT: Quade, Christian  
APPLICANT: Kruse, Fred  
APPLICANT: McGarity, Dennis  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM  
TITLE OF INVENTION: SECRETORY CELL LINES  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESS: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,582  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/028,427  
FILING DATE: 15-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/589,028  
FILING DATE: 19-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSD:514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear

US-08-784-582-10

Query Match 57.8%; Score 461; DB 3; Length 217;  
Best Local Similarity 70.3%; Pred. No. 7.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTIPLSRLFDVAMLRARLRHQLAFDYQEFEEAYIPKQKYSFLONPOTSLSSESISTPT 61  
DB 27 PPTIPLSRLFDVAMLRARLRHQLAFDYQEFEEAYIPKQKYSFLONPOTSLSSESISTPT 86  
QY 62 PSNRETOOKSNLELRISLLISLQSWLEPVOLGTGRFVNOHLGCS-----HLVE 111  
DB 87 PSNRETOOKSNLELRISLLISLQSWLEPVOLGTGRFVNOHLGCS-----HLVE 145  
QY 112 ALYVCG--ERGFYTPKTRGIYVQ 134  
DB 146 GIQTLMGRLDQ---SPRTGQIFRQ 167

US-08-785-271-10

RESULT 8  
Sequence 10, Application US/08785271  
Patent No. 6194176  
GENERAL INFORMATION:  
APPLICANT: Newgard, Christopher B.

APPLICANT: Halban, Philippe A.  
APPLICANT: No. 6194176/Minington, Karl D.  
APPLICANT: Clark, Samuel A.  
APPLICANT: Thigpen, Anice E.  
APPLICANT: Quade, Christian  
APPLICANT: Kruse, Fred  
APPLICANT: McGarity, Dennis  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM  
TITLE OF INVENTION: SECRETORY CELL LINES  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESS: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/785,271  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/589,028  
FILING DATE: 19-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSD:513  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear

US-08-785-271-10

Query Match 57.8%; Score 461; DB 3; Length 217;  
Best Local Similarity 70.3%; Pred. No. 7.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTIPLSRLFDVAMLRARLRHQLAFDYQEFEEAYIPKQKYSFLONPOTSLSSESISTPT 61  
DB 27 PPTIPLSRLFDVAMLRARLRHQLAFDYQEFEEAYIPKQKYSFLONPOTSLSSESISTPT 86  
QY 62 PSNRETOOKSNLELRISLLISLQSWLEPVOLGTGRFVNOHLGCS-----HLVE 111  
DB 87 PSNRETOOKSNLELRISLLISLQSWLEPVOLGTGRFVNOHLGCS-----HLVE 145  
QY 112 ALYVCG--ERGFYTPKTRGIYVQ 134  
DB 146 GIQTLMGRLDQ---SPRTGQIFRQ 167

US-08-759-628-11

RESULT 9  
Sequence 11, Application US/08759628  
Patent No. 6225446  
GENERAL INFORMATION:  
APPLICANT: Altman, Scott W.  
APPLICANT: Rock, Fernando L.  
APPLICANT: Bazan, J. Fernando  
APPLICANT: Kastelehn, Robert A.  
TITLE OF INVENTION: MUTATIONAL VARIANTS OF MAMMALIAN PROTEINS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: DNAX Research Institute

STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/759,628  
FILING DATE: 05-DEC-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,574  
FILING DATE: 06-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: DX0552Q  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-852-9196  
TELEFAX: 415-496-1200  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 32..53  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 94..115  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 133..153  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 192..210  
OTHER INFORMATION: /note="The peptides above are  
OTHER INFORMATION: depicted in Figure 1"  
US-08-759-628-11

Query Match 57.8%; Score 461; DB 3; Length 217;

Best Local Similarity 70.3%; Pred. No. 7.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLONPQTSLSFBSIPT 61  
DB 27 PPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLONPQTSLSFBSIPT 86  
QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQGTGPRFVNOHLCGS-----HIVE 111  
DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQGTGPRFVNOHLCGS-----HIVE 145  
QY 112 ALYIVCG--ERGFYTPKTRGIVEQ 134  
DB 146 GIQTLMGRLSDG--SPRTGQIFKQ 167  
RESULT 10  
US-09-284-878-1  
Sequence 1, Application US/09284878  
GENERAL INFORMATION:  
APPLICANT: Olazaran, Martha Guerrero  
APPLICANT: Saldaña, Hugo Barrera  
APPLICANT: Salgado, Jose Maria Vidar  
TITLE OF INVENTION: Genetically Modified Methylotrophic P. pastoris Yeast for the

TITLE OF INVENTION: Production and Secretion of the Human Growth Hormone  
FILE REFERENCE: 1829, 00100000  
CURRENT APPLICATION NUMBER: US/09/284,878  
CURRENT FILING DATE: 1999-07-21  
PRIOR APPLICATION NUMBER: PCT/MX97/00033  
PRIOR FILING DATE: 1997-10-24  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO: 1  
LENGTH: 217  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-284-878-1

Query Match 57.8%; Score 461; DB 3; Length 217;

Best Local Similarity 70.3%; Pred. No. 7.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLONPQTSLSFBSIPT 61  
DB 27 PPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLONPQTSLSFBSIPT 86  
QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQGTGPRFVNOHLCGS-----HIVE 111  
DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQGTGPRFVNOHLCGS-----HIVE 145  
QY 112 ALYIVCG--ERGFYTPKTRGIVEQ 134  
DB 146 GIQTLMGRLSDG--SPRTGQIFKQ 167

RESULT 11

US-09-929-918-9  
Sequence 9, Application US/09929918

PATENT NO. 6773899  
GENERAL INFORMATION:  
APPLICANT: Kordyum, Vitaliy A.  
APPLICANT: Chernykh, Svetlana I.  
APPLICANT: Slavchenko, Iryna Yu.  
APPLICANT: Vozianov, Oleksandr  
TITLE OF INVENTION: PHAGE-DEPENDENT SUPER PRODUCTION OF  
TITLE OF INVENTION: BIOLOGICALLY ACTIVE PROTEIN AND PEPTIDES  
FILE REFERENCE: PHAGE.006A  
CURRENT APPLICATION NUMBER: US/09/929,918  
CURRENT FILING DATE: 2001-08-15  
PRIOR APPLICATION NUMBER: 09/318,288  
PRIOR FILING DATE: 1999-05-25  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 9  
LENGTH: 217  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-929-918-9

Query Match 57.8%; Score 461; DB 4; Length 217;

Best Local Similarity 70.3%; Pred. No. 7.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLONPQTSLSFBSIPT 61  
DB 27 PPTPLSLRFDNMLRAHRLHQLAFTYQEFEEAYIPKQKYSFLONPQTSLSFBSIPT 86  
QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQGTGPRFVNOHLCGS-----HIVE 111  
DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQGTGPRFVNOHLCGS-----HIVE 145  
QY 112 ALYIVCG--ERGFYTPKTRGIVEQ 134  
DB 146 GIQTLMGRLSDG--SPRTGQIFKQ 167

RESULT 12

US-09-424-6208-25  
Sequence 25, Application US/09424620B  
Patent No. 6391585  
GENERAL INFORMATION:  
APPLICANT: HANIL SYNTHETIC FIBER CO., LTD.  
JANG, Ki-Ryong  
MOON, Jae-Moong  
BAE, Cheon-Soon  
YANG, Doo-Suk  
LEE, Jee-Mon  
TITLE OF INVENTION: Process for preparing recombinant proteins using highly efficient expression vector from Saccharomyces cerevisiae  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BACHMAN & LAPOINTE, P.C.  
STREET: Suite 1201, 900 Chapel Street  
CITY: New Haven  
STATE: Connecticut  
COUNTRY: U.S.A.  
ZIP: 06510-2802  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage  
COMPUTER: IBM  
OPERATING SYSTEM: WINDOWS 95/98  
SOFTWARE: MS WORD  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/424,620B  
FILING DATE: 24-NOV-6391585-1999  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 241 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: PROTEIN  
SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
US-09-424-6208-25  
Query Match 57.8%; Score 461; DB 3; Length 241;  
Best Local Similarity 70.3%; Pred. No. 8.2e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;  
QY 2 PPTPLSLRFDNMLRAHRLHQLAFDTYQSEFEAYIPKQKYSFLQNPQTSLSFSESIP 61  
DB 51 PPTPLSLRFDNMLRAHRLHQLAFDTYQSEFEAYIPKQKYSFLQNPQTSLSFSESIP 110  
QY 62 PSNRETOOKSNLELRISLLIQSWLEPVQF-LRSVFANSILVYGASDSNVDLKDLE 111  
DB 111 PSNRETOOKSNLELRISLLIQSWLEPVQF-LRSVFANSILVYGASDSNVDLKDLE 169  
QY 112 ALYVCG--ERGFYPTKRGIVEQ 134  
DB 170 GIOTMGRLDGG--SPRTGQIFKQ 191  
RESULT 13  
US-09-280-030-66  
Sequence 66, Application US/09280030A  
Patent No. 6506595  
GENERAL INFORMATION:  
APPLICANT: Sato, Seiji  
APPLICANT: Higashikuni, Naohiko  
APPLICANT: Kudo, Toshiyuki  
APPLICANT: Kondo, Masaaki  
TITLE OF INVENTION: DNA ENCODING NEW FUSION PROTEINS AND PROCESSES FOR THE  
TITLE OF INVENTION: PREPARING USEFUL POLYPEPTIDES THROUGH EXPRESSION OF THE  
FILE REFERENCE: 382,1026  
CURRENT APPLICATION NUMBER: US/09/280,030A  
CURRENT FILING DATE: 1999-03-26  
EARLIER APPLICATION NUMBER: JP10-87339/1998  
EARLIER FILING DATE: 1998-03-31  
NUMBER OF SEQ ID NOS: 66

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Designated is  
OTHER INFORMATION: an amino acid sequence of MWpSp-MWpM20-TEV-G-CH  
US-09-280-030-66  
Query Match 57.8%; Score 461; DB 4; Length 245;  
Best Local Similarity 70.3%; Pred. No. 8.4e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;  
QY 2 PPTPLSLRFDNMLRAHRLHQLAFDTYQSEFEAYIPKQKYSFLQNPQTSLSFSESIP 61  
DB 55 PPTPLSLRFDNMLRAHRLHQLAFDTYQSEFEAYIPKQKYSFLQNPQTSLSFSESIP 114  
QY 62 PSNRETOOKSNLELRISLLIQSWLEPVQF-LRSVFANSILVYGASDSNVDLKDLE 111  
DB 115 PSNRETOOKSNLELRISLLIQSWLEPVQF-LRSVFANSILVYGASDSNVDLKDLE 173  
QY 112 ALYVCG--ERGFYPTKRGIVEQ 134  
DB 174 GIOTMGRLDGG--SPRTGQIFKQ 195  
RESULT 14  
US-08-784-582-71  
Sequence 71, Application US/08784582  
Patent No. 6110707  
GENERAL INFORMATION:  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Halban, Philippe A.  
APPLICANT: No. 6110707Mingdon, Karl D.  
APPLICANT: Clark, Samuel A.  
APPLICANT: Thigpen, Antice E.  
APPLICANT: Quade, Christian  
APPLICANT: Kruse, Fred  
APPLICANT: Mcgarity, Dennis  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM  
TITLE OF INVENTION: SECRETORY CELL LINES  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,582  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/028,427  
FILING DATE: 15-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/589,028  
FILING DATE: 19-04N-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSD,514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 71:

SEQUENCE CHARACTERISTICS:  
LENGTH: 274 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-784-582-71

Query Match  
Best Local Similarity 70.3%; Pred. No. 9.7e-42;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTPLSLRPLFNAMLRARHLQALAFDTYQEFEEAYIPKEOKYSEFLQNPQTSLSFSSES IPT 61  
DB 27 FFTPLSLRPLFNAMLRARHLQALAFDTYQEFEEAYIPKEOKYSEFLQNPQTSLSFSSES IPT 86  
QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQLGTGPRFVNQHLCS-----HIVE 111  
DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDL EE 145  
QY 112 ALIYVCG--ERGFYTPKTRGIVEQ 134  
DB 146 GIQTLGRLEDG---SPRTQIFKQ 167

RESULT 15  
US-08-784-582-73  
Sequence 73, Application US/08784582  
Patent No. 6110707  
GENERAL INFORMATION:  
APPLICANT: Newgard, Christopher B.  
APPLICANT: Halban, Philippe A.  
APPLICANT: No. 6110707/England, Karl D.  
APPLICANT: Clark, Samuel A.  
APPLICANT: Thigpen, Amice E.  
APPLICANT: Quade, Christian  
APPLICANT: Kruse, Fred  
APPLICANT: McGarry, Dennis  
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF PROTEINS FROM  
NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/784,582  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/028,427  
FILING DATE: 15-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/589,028  
FILING DATE: 19-JAN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Highlander, Steven L.  
REGISTRATION NUMBER: 37,642  
REFERENCE/DOCKET NUMBER: UTSD:514  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 73:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 360 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear  
US-08-784-582-73

Query Match  
Best Local Similarity 70.3%; Pred. No. 1.4e-41;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 FFTPLSLRPLFNAMLRARHLQALAFDTYQEFEEAYIPKEOKYSEFLQNPQTSLSFSSES IPT 61  
DB 27 FFTPLSLRPLFNAMLRARHLQALAFDTYQEFEEAYIPKEOKYSEFLQNPQTSLSFSSES IPT 86  
QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQLGTGPRFVNQHLCS-----HIVE 111  
DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQF-LRSVFANSLVYGASDSNVYDLKDL EE 145  
QY 112 ALIYVCG--ERGFYTPKTRGIVEQ 134  
DB 146 GIQTLGRLEDG---SPRTQIFKQ 167

Search completed: November 2, 2004, 20:24:36  
Job time : 37.5314 secs

**This Page Blank (usp10)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 2, 2004, 19:59:41; Search time 27.6753 Seconds  
(without alignments)  
521,495 Million cell updates/sec

Title: US-10-054-873-7

Perfect score: 797  
Sequence: 1 MPTPIPLSRLEFDNMLRAHR.....IVQCCTCSICSLYQLENYCN 150

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database:

1: PIR79:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	461	57.8	217	1	STHU
2	460	57.7	217	2	167410
3	426.5	53.5	217	1	STHUV
4	426.5	53.5	256	1	STHUV2
5	407.5	51.1	217	2	167409
6	405	50.8	217	2	167411
7	396	49.7	212	2	167408
8	396	49.7	212	2	153267
9	381	47.8	217	1	167410
10	381	47.8	217	1	167410
11	359.5	46.1	215	2	167410
12	310.5	33.0	216	2	167410
13	307.5	33.6	190	2	167410
14	306.5	33.6	190	2	167410
15	304.5	33.2	216	1	167410
16	302.5	33.0	216	1	167410
17	302.5	33.0	216	1	167410
18	301.5	32.8	190	2	167410
19	301.5	32.8	190	2	167410
20	301.5	32.8	216	1	167410
21	301.5	32.8	216	1	167410
22	299.5	32.8	216	1	167410
23	297.5	32.8	190	1	167410
24	295.5	32.8	190	1	167410
25	289.5	32.8	217	1	167410
26	289.5	32.8	217	1	167410
27	289.5	32.8	217	1	167410
28	289.5	32.8	217	1	167410
29	278.5	34.9	216	2	167410

30	277.5	34.8	110	1	INRB
31	277.5	34.8	110	2	B42179
32	275.5	34.6	216	2	A60509
33	275	34.5	96	2	PC7082
34	273.5	34.3	51	1	INEL
35	273.5	34.3	51	1	INEL
36	273.5	34.3	51	1	INEL
37	273.5	34.3	110	2	UQ0178
38	272	34.1	110	2	A42179
39	271.5	34.1	51	1	INHY
40	270	33.9	110	1	IPHU
41	268.5	33.7	51	1	INMSP
42	268.5	33.7	191	2	A60625
43	267.5	33.6	51	2	A59151
44	266.5	33.4	105	2	IPBO
45	265.5	33.3	110	2	I48166

## ALIGNMENTS

## RESULT 1

STHU

somatotropin 1 precursor [validated] - human

N/Alternate names: growth hormone 1; hGH-N; pituitary somatotropin

N/Contains: growth hormone 5K peptide; somatotropin 1, long form; somatotropin 1, short

C/Species: Homo sapiens (man)

C/Date: 24-Apr-1984 #sequence revision 10-Feb-1995 #ext change 09-Jul-2004

C/Accession: A93731; A32435; A36594; A94247; A90051; A93197; A93778; A91664; A90217; A

R/Denoto, F.M.; Moore, D.D.; Goodman, H.M.

Nucleic Acids Res. 9, 3719-3730, 1981

A/Title: Human growth hormone DNA sequence and mRNA structure: possible alternative sp

A/Reference number: A93731; MUID:82014939; PMID:6269091

A/Accession: A93731

A/Molecule type: DNA

A/Residues: 1-217 <DEN>

A/Cross-references: UNIPROT:P01241; GB:V00520

A/Note: the 20K short form somatotropin lacks residues 58-72 (32-46 in the active horm

R/Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seeburg,

Genomics 4, 479-497, 1989

A/Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.

A/Reference number: A32435; MUID:89307277; PMID:2744760

A/Accession: A32435

A/Molecule type: mRNA

A/Residues: 1-217 <ROS>

A/Cross-references: GB:V00519

A/Note: 35-Pro was also found

R/Martell, U.A.; Hallowell, R.A.; Baxter, J.D.; Goodman, H.M.

Science 205, 602-607, 1979

A/Title: Human growth hormone: complementary DNA cloning and expression in bacteria.

A/Reference number: A94247; MUID:79203293; PMID:77496

A/Accession: A94247

A/Molecule type: mRNA

A/Residues: 1-217 <MAR>

R/Li, C.H.; Dixon, J.S.; Liu, W.K.

Arch. Biochem. Biophys. 133, 70-91, 1969

A/Title: Human pituitary growth hormone. XIX. The primary structure of the hormone.

A/Reference number: A90048; MUID:69289202; PMID:5810834

A/Contents: annotation

R/Li, C.H.; Dixon, J.S.

Arch. Biochem. Biophys. 146, 233-236, 1971

A/Title: Human pituitary growth hormone. XXXII. The primary structure of the hormone: x

A/Reference number: A90051; MUID:7213935; PMID:5144027

A/Accession: A90051

A/Molecule type: protein

A:Residues: 27-94;96-217 <LIC>  
 R:Niall, H.D.  
 Nature New Biol. 230, 90-91, 1971  
 A:Title: Revised primary structure for human growth hormone.  
 A:Reference number: A93397; MUID:71139765; PMID:5279046  
 A:Accession: A93397  
 A:Molecule type: protein  
 A:Residues: 27-51 <NIA>  
 R:Niall, H.D.; Hogan, M.L.; Sauer, R.; Rosenblum, I.V.; Greenwood, F.C.  
 Proc. Natl. Acad. Sci. U.S.A. 68, 866-869, 1971  
 A:Title: Sequences of pituitary and placental lactogenic and growth hormones: evolution  
 A:Reference number: A93778; MUID:71153968; PMID:5279528  
 A:Accession: A93778  
 A:Molecule type: protein  
 A:Residues: 119-120;157-159 <NI2>  
 R:Niall, H.D.  
 in Prolactin and Carcinogenesis, Proc. Fourth Tenoux Workshop Prolactin, Griffiths, K.,  
 A:Title: The chemistry of the human lactogenic hormones.  
 A:Reference number: A94427  
 A:Accession: A94427  
 A:Contents: annotation; somatotropin revision  
 R:Bewley, T.A.; Dixon, J.S.; Li, C.H.  
 Int. J. Pept. Protein Res. 4, 281-287, 1972  
 A:Title: Sequence comparison of human pituitary growth hormone, human chorionic somatoma  
 A:Reference number: A91764; MUID:73092028; PMID:4675454  
 A:Accession: A91764  
 A:Molecule type: protein  
 A:Residues: 27-217 <BBW>  
 R:Lewis, U.J.; Bonewald, L.F.; Lewis, L.J.  
 Biochem. Biophys. Res. Commun. 92, 511-516, 1980  
 A:Title: The 20,000-dalton variant of human growth hormone: location of the amino acid  
 A:Reference number: A90217; MUID:80130196; PMID:7356479  
 A:Accession: A90217  
 A:Contents: somatotropin, 20K short variant  
 A:Molecule type: protein  
 A:Residues: 46-57;73-80 <LEW>  
 R:Chapman, G.E.; Rogers, K.M.; Brittain, T.; Bradshaw, R.A.; Bates, O.J.; Turner, C.; Ca  
 J. Biol. Chem. 256, 2395-2401, 1981  
 A:Title: The 20,000 molecular weight variant of human growth hormone. Preparation and se  
 A:Reference number: A92311; MUID:8117361; PMID:7462247  
 A:Accession: A92311  
 A:Molecule type: protein  
 A:Residues: 27-57;73-79 <CHA>  
 R:Singh, R.N.P.; Seavey, B.K.; Lewis, L.J.; Lewis, U.J.  
 J. Protein Chem. 2, 425-436, 1983  
 A:Title: Human growth hormone peptide 1-43: isolation from pituitary glands.  
 A:Reference number: A61466  
 A:Accession: A61466  
 A:Molecule type: protein  
 A:Residues: 27-69 <SIN>  
 A:Note: growth hormone 5K peptide has insulin potentiating activity; its physiological f  
 R:Robson, V.M.J.; Rae, I.D.; NG, F.  
 Biol. Chem. Hoppe-Seyler 371, 423-431, 1990  
 A:Title: Identification of the aspartamide structure in a previously-reported peptide.  
 A:Reference number: S09685; MUID:9034745; PMID:2378679  
 A:Accession: S09685  
 A:Molecule type: protein  
 A:Residues: 27-34; 'L', '36-47 <ROB>  
 R:de Vos, A.M.; Ultsch, M.; Kossiakoff, A.A.  
 Science 255, 306-312, 1992  
 A:Title: Human growth hormone and extracellular domain of its receptor: crystal structu  
 A:Reference number: A41728; MUID:92196577; PMID:1549776  
 A:Accession: A41728  
 A:Contents: annotation; X-ray crystallography, 2.8 angstroms  
 A:Note: the structure of the complex with growth hormone receptor is described  
 R:Gray, G.L.; Balridge, J.S.; McKeown, K.S.; Heyneker, H.L.; Chang, C.N.  
 Gene 39, 247-254, 1985  
 A:Title: Periplasmic production of correctly processed human growth hormone in Escherich  
 A:Reference number: 141126; MUID:86137393; PMID:3312261  
 A:Accession: 184549  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-26 <RES>  
 A:Cross-references: GB:M14398; NID:9183158; PIDN:AAAS2554.1; PID:9183159

C:Comment: The gene for this hormone is transcribed only in somatotrophic cells of the  
 C:Comment: About 90% of somatotropin is the 22K long form.  
 C:Genetics:  
 A:Gene: GDB:GH1  
 A:Cross-references: GDB:119962; OMIM:139250  
 A:Map position: 17q23.1-17q23.3  
 A:Introns: 4/1; 57/3; 97/3; 152/3  
 C:Superfamily: prolactin  
 C:Keywords: alternative splicing; hormone; pituitary  
 F:1-26/Domain: signal sequence #status predicted <SIG>  
 F:27-217/Product: somatotropin 1, long form #status experimental <SOL>  
 F:27-69/Product: growth hormone 5K peptide #status experimental <SKP>  
 F:27-57;73-217/Product: somatotropin 1, short form #status experimental <SOS>  
 F:79-191;208-215/Diulfide bonds: #status experimental

Query Match 57.8%; Score 461; DB 1; Length 217;  
 Best Local Similarity 70.3%; Pred. No. 2.8e-37;  
 Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PFTPLSLFNDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSISPT 61  
 DB 27 PFTPLSLFNDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSISPT 86  
 QY 62 PSNRRETOCKSNLELRISLLIQSWLEPVLGSPFVNOHCGS-----HYVE 111  
 DB 87 PSNRRETOCKSNLELRISLLIQSWLEPVLGSPFVNOHCGS-----HYVE 145  
 QY 112 ALYLVCG--ERGFYTPKTRGIVDQ 134  
 DB 146 GIQLTWGRLEDS--SPRTGQIFRQ 167

RESULT 2  
 167410  
 somatotropin - rhesus macaque  
 N:Alternate names: growth hormone  
 C:Species: Macaca mulatta (rhesus macaque)  
 C:Date: 31-May-1996 #sequence revision 31-May-1996 #text change 09-Jul-2004  
 A:Accession: 167410; A05094  
 R:Gollos, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.  
 Endocrinology 133, 1744-1752, 1993  
 A:Title: Cloning of four growth hormone/chorionic somatomotropin-related compleme  
 A:Reference number: 153267; MUID:94008724; PMID:8404617  
 A:Accession: 167410  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-217 <RES>  
 A:Cross-references: UNIPROT:P33093; GB:L16556; NID:9293114; PIDN:AAAI8842.1; PID:929311  
 R:Li, C.H.; Chung, D.; Lam, H.W.; Stein, S.  
 Arch. Biochem. Biophys. 245, 287-291, 1986  
 A:Title: The primary structure of monkey pituitary growth hormone.  
 A:Reference number: A05094; MUID:86129460; PMID:3060959  
 A:Accession: A05094  
 A:Molecule type: protein  
 A:Residues: 27-99; 'Q', '101-178; 'D', '180-217 <LIC>  
 A:Note: the monkey species is not identified in the reference  
 R:Raben, M.S.  
 Science 125, 883-884, 1957  
 A:Title: Preparation of growth hormone from pituitaries of man and monkey.  
 A:Reference number: A44774  
 A:Accession: A44774  
 A:Contents: annotation; identification of source organism  
 C:Superfamily: prolactin

Query Match 57.7%; Score 460; DB 2; Length 217;  
 Best Local Similarity 98.9%; Pred. No. 3.5e-37;  
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 PFTPLSLFNDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSISPT 61  
 DB 27 PFTPLSLFNDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSISPT 86  
 QY 62 PSNRRETOCKSNLELRISLLIQSWLEPVLGSPFVNOHCGS-----HYVE 111  
 DB 87 PSNRRETOCKSNLELRISLLIQSWLEPVLGSPFVNOHCGS-----HYVE 145

DB 87 PSNRETOOKSNLELRLISILLIOSWLEPVQ 117

## RESULT 3

STHUV

somatotropin 2 precursor - human

N.Alternate names: growth hormone 2; growth hormone variant; hGH-V; placental somatotropin

C.Species: Homo sapiens (man)

C.Date: 17-Dec-1983 #sequence\_revision 10-Feb-1995 #text\_change 09-Jul-2004

C.Accession: D32435; B28072; A01511; I52104; A60711

R.Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Salana, H.A.; Gellinas, R.E.; Seeburg, P.

A.Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.

A.Reference number: A32435; MUID:89307277; PMID:2744760

A.Accession: D32435

A.Molecule type: DNA

A.Residues: 1-217 &lt;CHE&gt;

A.Cross-references: UNIPROT:P01242; GB:J03071; NID:G183148; PIND:AAA52552.1; PID:G183152

J.Cooke, N.E.; Ray, J.; Emery, J.G.; Liebhauer, S.A.

A.Title: Two distinct species of human growth hormone-variant mRNA in the human placenta

A.Reference number: A92725; MUID:88243769; PMID:3379057

A.Accession: B28072

A.Molecule type: mRNA

A.Residues: 1-217 &lt;COO&gt;

R.Seeburg, P.H.

DNA 1, 239-249, 1982

A.Title: The human growth hormone gene family: nucleotide sequences show recent divergen

A.Reference number: A01511; MUID:83182010; PMID:7169009

A.Accession: A01511

A.Molecule type: DNA

A.Residues: 1-34, 'P', 36-217 &lt;SEE&gt;

R.Ignotz, A.; Scippo, M.L.; Franken, F.; Hennen, G.

Arch. Int. Physiol. Biochim. 96, 63-67, 1988

A.Title: Cloning and nucleotide sequence of placental hGH-V cDNA.

A.Reference number: I52104; MUID:89024984; PMID:2160050

A.Accession: I52104

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-217 &lt;IGO&gt;

R.Franken, F.; Scippo, M.L.; Van Beeumen, J.; Ignotz, A.; Hennen, G.

J.Clin. Endocrinol. Metab. 71, 15-18, 1990

A.Title: Identification of placental human growth hormone as the growth hormone-V gene

A.Reference number: A60711; MUID:90317018; PMID:2156278

A.Accession: A60711

A.Molecule type: protein

A.Residues: 27-44; 46-57 &lt;FRA&gt;

A.Experimental source: tissue placenta

C.Comment: This gene is expressed by the placenta.

C.Genetics:

A.Gene: GDB:GH2

A.Cross-references: GDB:119983; OMIM:139240

A.Map position: 17q22-17q24

A.Introns: 4/1; 57/3; 97/3; 152/3

C.Superfamily: prolactin

C.Keywords: alternative splicing; glycoprotein; hormone; placenta

F.1-26/Domains: signal sequence #status predicted &lt;SIG&gt;

F.27-217/Product: somatotropin 2, long splice form #status predicted &lt;SOL&gt;

F.127-57, 73-217/Product: somatotropin 2, short splice form #status predicted &lt;SOS&gt;

F.79-191, 208-215/Dissulfide bonds: #status predicted

F.166/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match

Best Local Similarity 53.5%; Score 426.5; DB 1; Length 217;

Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;

DB 27 PPTPLSLRFDNMLRAHRLHQLAFDPTYOEFEEAYIPKEQKYSFLONPQTSLSFSSESIP 61

PPTPLSLRFDNMLRAHRLHQLAFDPTYOEFEEAYIPKEQKYSFLONPQTSLSFSSESIP 86

QY 62 PSNRETOOKSNLELRLISILLIOSWLEPVQ-----GTGPRFVNQHL 104

DB 87 PSNRETOOKSNLELRLISILLIOSWLEPVQLRSVPANSLVYGADSNVYRL 140

## RESULT 4

STHUV2

somatotropin 2 precursor, splice form 2 - human

N.Alternate names: growth hormone variant-2; placental somatotropin form 2

C.Species: Homo sapiens (man)

C.Date: 30-Sep-1989 #sequence\_revision 10-Feb-1995 #text\_change 09-Jul-2004

C.Accession: A28072

R.Cooke, N.E.; Ray, J.; Emery, J.G.; Liebhauer, S.A.

J. Biol. Chem. 263, 9001-9006, 1988

A.Title: Two distinct species of human growth hormone-variant mRNA in the human placenta

A.Reference number: A92725; MUID:88243769; PMID:3379057

A.Accession: A28072

A.Molecule type: mRNA

A.Residues: 1-256 &lt;COO&gt;

A.Cross-references: UNIPROT:P01242

A.Note: an alternative splice junction for intron 4 is used

C.Genetics:

A.Gene: GDB:GH2

A.Cross-references: GDB:119983; OMIM:139240

A.Map position: 17q22-17q24

A.Introns: 4/1; 57/3; 97/3; 152/3

C.Superfamily: prolactin

C.Keywords: alternative splicing; hormone; placenta

F.1-26/Domains: signal sequence #status predicted &lt;SIG&gt;

F.27-256/Product: somatotropin 2 splice form 2 #status predicted &lt;WAT&gt;

Query Match

Best Local Similarity 53.5%; Score 426.5; DB 1; Length 256;

Matches 89; Conservative 4; Mismatches 10; Indels 11; Gaps 1;

DB 2 PPTPLSLRFDNMLRAHRLHQLAFDPTYOEFEEAYIPKEQKYSFLONPQTSLSFSSESIP 61

PPTPLSLRFDNMLRAHRLHQLAFDPTYOEFEEAYIPKEQKYSFLONPQTSLSFSSESIP 86

QY 62 PSNRETOOKSNLELRLISILLIOSWLEPVQ-----GTGPRFVNQHL 104

DB 87 PSNRETOOKSNLELRLISILLIOSWLEPVQLRSVPANSLVYGADSNVYRL 140

## RESULT 5

chorionic somatomammotropin-3 - rhesus macaque

C.Species: Macaca mulatta (rhesus macaque)

C.Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004

C.Accession: I67409

R.Golos, T.G.; Durning, M.; Fisher, J.M.; Towler, P.D.

Endocrinology 133, 1744-1752, 1993

A.Title: Cloning of four growth hormone/chorionic somatomammotropin-related complements

A.Reference number: I53267; MUID:94008724; PMID:8404617

A.Accession: I67409

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: mRNA

A.Residues: 1-217 &lt;RES&gt;

A.Cross-references: UNIPROT:Q07369; GB:U16554; NID:G293112; PIND:AAA18641.1; PID:G29311

C.Superfamily: prolactin

Query Match

Best Local Similarity 51.1%; Score 407.5; DB 2; Length 217;

Matches 82; Conservative 12; Mismatches 19; Indels 1; Gaps 1;

DB 3 PPTPLSLRFDNMLRAHRLHQLAFDPTYOEFEEAYIPKEQKYSFLONPQTSLSFSSESIP 62

PPTPLSLRFDNMLRAHRLHQLAFDPTYOEFEEAYIPKEQKYSFLONPQTSLSFSSESIP 87

QY 63 SNRETOOKSNLELRLISILLIOSWLEPVQLTGPRFVNQHLGSHLVEALYIV 116

DB 88 SNRETOOKSNLELRLISILLIOSWLEPVQLTGVFANLVYGTSDAYDL 140

## RESULT 6

157411

Somatotropin - rhesus macaque

N/Alternate names: growth hormone

C/Species: Macaca mulatta (rhesus macaque)

C/Date: 31-May-1996 #sequence\_revision 31-May-1996 #ext\_change 09-Jul-2004

C/Accession: 157411

R/Golios, T.G.; Durning, M.; Fisher, J.M.; Fowler, P.D.

Endocrinology 133, 1744-1752, 1993

A/Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementat

A/Reference number: 153267; MUID:94008724; PMID:8404617

A/Accession: 157411

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-217 &lt;RES&gt;

A/Cross-references: UNIPROT:Q07370; GB:U16555; NID:g293116; PIDN:AAA0180.1; PID:g293117

C/Superfamily: prolactin

Query Match

Best Local Similarity

Matches

86; Conservative

9; Mismatches

18; Indels

14; Gaps

2;

50.8%; Score 405; DB 2; Length 217;

67.7%; Pred. No. 7.6e-32;

6; Mismatches

18; Indels

14; Gaps

2;

50.8%; Score 405; DB 2; Length 217;

67.7%; Pred. No. 7.6e-32;

6; Mismatches

18; Indels

14; Gaps

2;

50.8%; Score 405; DB 2; Length 217;

67.7%; Pred. No. 7.6e-32;

6; Mismatches

18; Indels

14; Gaps

2;

50.8%; Score 405; DB 2; Length 217;

67.7%; Pred. No. 7.6e-32;

6; Mismatches

18; Indels

14; Gaps

2;

50.8%; Score 405; DB 2; Length 217;

67.7%; Pred. No. 7.6e-32;

6; Mismatches

18; Indels

14; Gaps

2;

Endocrinology 133, 1744-1752, 1993

A/Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementat

A/Reference number: 153267; MUID:94008724; PMID:8404617

A/Accession: 153267

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-217 &lt;RES&gt;

A/Cross-references: UNIPROT:Q07367; GB:U16552; NID:g293108; PIDN:AAA1883.1; PID:g293109

C/Superfamily: prolactin

Query Match

Best Local Similarity

Matches

74; Conservative

11; Mismatches

5; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

Endocrinology 133, 1744-1752, 1993

A/Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementat

A/Reference number: 153267; MUID:94008724; PMID:8404617

A/Accession: 153267

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-217 &lt;RES&gt;

A/Cross-references: UNIPROT:Q07367; GB:U16552; NID:g293108; PIDN:AAA1883.1; PID:g293109

C/Superfamily: prolactin

Query Match

Best Local Similarity

Matches

74; Conservative

11; Mismatches

5; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

Endocrinology 133, 1744-1752, 1993

A/Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementat

A/Reference number: 153267; MUID:94008724; PMID:8404617

A/Accession: 153267

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-217 &lt;RES&gt;

A/Cross-references: UNIPROT:Q07367; GB:U16552; NID:g293108; PIDN:AAA1883.1; PID:g293109

C/Superfamily: prolactin

Query Match

Best Local Similarity

Matches

74; Conservative

11; Mismatches

5; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

Endocrinology 133, 1744-1752, 1993

A/Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementat

A/Reference number: 153267; MUID:94008724; PMID:8404617

A/Accession: 153267

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-217 &lt;RES&gt;

A/Cross-references: UNIPROT:Q07367; GB:U16552; NID:g293108; PIDN:AAA1883.1; PID:g293109

C/Superfamily: prolactin

Query Match

Best Local Similarity

Matches

74; Conservative

11; Mismatches

5; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

Endocrinology 133, 1744-1752, 1993

A/Title: Cloning of four growth hormone/chorionic somatomammotropin-related complementat

A/Reference number: 153267; MUID:94008724; PMID:8404617

A/Accession: 153267

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-217 &lt;RES&gt;

A/Cross-references: UNIPROT:Q07367; GB:U16552; NID:g293108; PIDN:AAA1883.1; PID:g293109

C/Superfamily: prolactin

Query Match

Best Local Similarity

Matches

74; Conservative

11; Mismatches

5; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

0;

49.7%; Score 396; DB 2; Length 217;

82.2%; Pred. No. 5.7e-31;

5; Mismatches

11; Indels

0; Gaps

A:Accession: A90054  
 A:Molecule type: protein  
 A:Residues: 27-217 <LIC>  
 A:Experimental source: placenta  
 R:Wali, H.D.  
 In prolactin and Carcinogenesis, Proc. Fourth Tumor Workshop Prolactin, Griffiths, K.,  
 A:Title: The chemistry of the human lactogenic hormones.  
 A:Reference number: A94427  
 A:Accession: A94427  
 A:Molecule type: protein  
 A:Residues: 27-217 <NTA>  
 A:Experimental source: placenta  
 R:McC A Bhaidd, N.; Tibton, K.F.  
 Biochem. Soc. Trans. 19, 205, 1991  
 A:Title: Catechol-O-methyltransferase from human placenta: purification and some properties  
 A:Reference number: A61283; MUID:91244006; PMID:2037148  
 A:Accession: A61283  
 A:Molecule type: protein  
 A:Residues: 27-46 <NIC>  
 A:Note: chorionamniotopin apparently copurified with placental catechol-O-methyltransferase  
 R:Sherwood, L.M.; Handwerker, S.; McLaurin, W.D.; Lanner, M.  
 Nature New Biol. 233, 59-61, 1971  
 A:Title: Amino-acid sequence of human placental lactogen.  
 A:Reference number: A93401; MUID:72016313; PMID:5286363  
 A:Contents: annotation  
 R:Sherwood, L.M.; Handwerker, S.; McLaurin, W.D.; Lanner, M.  
 Nature New Biol. 233, 64, 1972  
 A:Reference number: A93405  
 A:Contents: annotation  
 R:Schneider, A.B.; Kowaleki, K.; Russell, J.; Sherwood, L.M.  
 J. Biol. Chem. 254, 3782-3787, 1979  
 A:Title: Identification of the interchain disulfide bonds of dimeric human placental lac  
 A:Reference number: A92251; MUID:79173081; PMID:438159  
 A:Contents: annotation; dimeric disulfide bonds  
 R:Selby, M.J.; Barra, A.; Baxter, J.D.; Bell, G.I.; Eberhardt, N.L.  
 J. Biol. Chem. 259, 13131-13138, 1984  
 A:Title: Analysis of a major human chorionic somatomammotropin gene. Evidence for two fu  
 A:Reference number: 155229; MUID:85030426; PMID:6208192  
 A:Accession: 155229  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-217 <RES>  
 A:Cross-references: GB:X02401; NID:9181120; PIDN:AAA52115.1; PID:9181121  
 R:Seeburg, P.H.; Shine, J.; Martial, J.A.; Ullrich, A.; Goodman, H.  
 Trans. Assoc. Am. Physicins 90, 109-116, 1977  
 A:Title: Nucleotide sequence of a human gene coding for a polypeptide hormone.  
 A:Accession: 159658; MUID:79160787; PMID:611657  
 A:Reference number: 159658  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 160-217 <RE2>  
 A:Cross-references: GB:M25118; NID:9181124; PIDN:AAA35721.1; PID:9181125  
 C:Genetics:  
 A:Gene: GDB:CSH1  
 A:Cross-references: GDB:119084; OMIM:150200  
 A:Map position: 17q22-17q24  
 A:Introns: 4/1; 57/3; 97/3; 152/3  
 C:Superfamily: prolactin  
 C:Keywords: hormone; placenta  
 F:1-26/DNA: signal sequence #status experimental <SIG>  
 F:27-217/Product: chorionamniotopin A #status experimental <MAT>  
 F:79-191/Disulfide bonds: #status experimental  
 F:208-215/Disulfide bonds: (in monomeric form) #status experimental  
 F:208-215/Disulfide bonds: interchain (to 215 in dimeric form) #status experimental  
 F:215/Disulfide bonds: interchain (to 208 in dimeric form) #status experimental

Query Match 47.8%; Score 381; DB 1; Length 217;  
 Best Local Similarity 82.0%; Pred. No. 1.6e-29;  
 Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 TTPSRFLPDNAMLRAHRLHQLAFTYQEFEEAYIPKQKXSFLONPQTSLSFSESIPTPS 63  
 DB 29 TVPLSRFLPDHMLQAHRAHQLAIDTYQEFEEYIPKQKXSFLONPQTSLSFSESIPTPS 88

QY 64 NREETQOKSNLELRISLLILIOGWLPEVQ 92  
 DB 89 NREETQOKSNLELRISLLILIOGWLPEVR 117

RESULT 10  
 E32435  
 Chorionamniotopin B precursor - human  
 N:Alternate names: chorionic somatomammotropin 2  
 C:Species: Homo sapiens (man)  
 C:Date: 29-Dec-1989 #sequence\_revision 29-Dec-1989 #text\_change 09-Jul-2004  
 C:Accession: E32435  
 R:Chen, E.Y.; Liao, Y.C.; Smith, D.H.; Barrera-Saldana, H.A.; Gellinas, R.E.; Seeburg, G.  
 Genomics 4, 479-497, 1989  
 A:Title: The human growth hormone locus: nucleotide sequence, biology, and evolution.  
 A:Reference number: A32435; MUID:8930727; PMID:2744760  
 A:Accession: E32435  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-217 <CHE>  
 A:Cross-references: UNIPROT:Q14407; GB:J03071; NID:9183148; PIDN:AAA52553.1; PID:91831  
 C:Genetics:  
 A:Gene: GDB:CSH2  
 A:Cross-references: GDB:119813; OMIM:118820  
 A:Map position: 17q22-17q24  
 C:Superfamily: prolactin

Query Match 47.8%; Score 381; DB 2; Length 217;  
 Best Local Similarity 82.0%; Pred. No. 1.6e-29;  
 Matches 73; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 4 TTPSRFLPDNAMLRAHRLHQLAFTYQEFEEAYIPKQKXSFLONPQTSLSFSESIPTPS 63  
 DB 29 TVPLSRFLPDHMLQAHRAHQLAIDTYQEFEEYIPKQKXSFLONPQTSLSFSESIPTPS 88

QY 64 NREETQOKSNLELRISLLILIOGWLPEVQ 92  
 DB 89 NREETQOKSNLELRISLLILIOGWLPEVR 117

RESULT 11  
 A26449  
 Chorionamniotopin precursor (allele hCS-3) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
 C:Accession: A26449  
 R:Hirt, H.; Krimmelman, J.; Birnbaum, M.J.; Chen, E.Y.; Seeburg, P.H.; Eberhardt, N.L.; E  
 DNA 6, 59-70, 1987  
 A:Title: The human growth hormone gene locus: structure, evolution, and allelic variat  
 A:Reference number: A26449; MUID:87161235; PMID:3030680  
 A:Accession: A26449  
 A:Molecule type: DNA  
 A:Residues: 1-215 <HIR>  
 A:Cross-references: UNIPROT:P01243  
 C:Superfamily: prolactin  
 F:1-26/DNA: signal sequence #status predicted <SIG>  
 F:27-215/Product: chorionamniotopin, hCS-3 allele #status predicted <MAT>

Query Match 45.1%; Score 359.5; DB 2; Length 215;  
 Best Local Similarity 80.5%; Pred. No. 1.9e-27;  
 Matches 70; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

QY 4 TTPSRFLPDNAMLRAHRLHQLAFTYQEFEEAYIPKQKXSFLONPQTSLSFSESIPTPS 63  
 DB 29 TVPLSRFLPDHMLQAHRAHQLAIDTYQEFEEYIPKQKXSFLONPQTSLSFSESIPTPS 88

QY 64 NREETQOKSNLELRISLLILIOGWLPEVQ 90  
 DB 89 NREETQOKSNLELRISLLILIOGWLPEVR 114

RESULT 12

B49159

somatotropin - golden hamster

N:Alternate names: growth hormone

C:Species: Mesocricetus auratus (golden hamster)

C&gt;Date: 19-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: B49159

R:Southard, J.N.; Sanchez-Dimenez, F.; Campbell, G.T.; Talamantes, F.

R:Endocrinology 129, 2965-2971, 1991

A:Title: Sequence and expression of hamster prolactin and growth hormone messenger RNAs.

A:Reference number: A49159; MUID:92063850; PMID:1954881

A:Accession: B49159

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-216 &lt;SOU&gt;

A:Cross-references: UNIPROT:P37886; GB:S66299; NID:G239355; PIDD:AA820368.1; PID:G239356

A&gt;Note: sequence extracted from NCBI backbone (NCBIN:66299, NCBIIP:66300)

C:Superfamily: prolactin

Query Match

Best Local Similarity 39.0%; Score 310.5; DB 2; Length 216;

Matches 61; Conservative 13; Mismatches 16; Indels 1; Gaps 1;

Db

2 FPTPLSLRFDNAMLRAHRLHQAFTYQEFEEAYIPKQKYSFLQNPQTSLSFSES IPT 61

27 FPMPLSSLSFANAVLRAQHQLADTYKEFERAYIPBQGRYS-IQNAQAFCSSETIPA 85

Qy

62 PSNRETOOKSNLELRISLLIQSWLEPVQ 92

86 PTGKEAQQRSDWELRLFSLLIQSWLGPVQ 116

Db

RESULT 13

PN0140

somatotropin - sei whale

N:Alternate names: growth hormone

C:Species: Balaenoptera borealis (sei whale)

C&gt;Date: 07-May-1993 #sequence\_revision 07-May-1993 #text\_change 09-Jul-2004

C:Accession: PN0140

R:Yudaev, N.A.; Pankov, Y.A.; Bulatov, A.A.; Osipova, T.A.

R:Biokhimiya 47, 1059-1069, 1982

A:Title: Amino acid sequence of sei whale somatotropin.

A:Reference number: PN0140; MUID:83000569; PMID:7115813

A:Accession: PN0140

A:Molecule type: protein

A:Residues: 1-190 &lt;YUD&gt;

A:Cross-references: UNIPROT:P33092

A&gt;Note: article in Russian with English abstract

C:Superfamily: prolactin

C:Keywords: growth factor; hormone

F:52-163,180-188/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 38.6%; Score 307.5; DB 2; Length 190;

Matches 61; Conservative 14; Mismatches 15; Indels 1; Gaps 1;

Db

2 FPTPLSLRFDNAMLRAHRLHQAFTYQEFEEAYIPKQKYSFLQNPQTSLSFSES IPT 61

1 FPMPLSSLSFANAVLRAQHQLADTYKEFERAYIPBQGRYS-FLQNAQSTGCFSEV IPT 59

Qy

62 PSNRETOOKSNLELRISLLIQSWLEPVQ 92

60 PAKDDEAQQRSDWELRLFSLLIQSWLGPVQ 90

Db

RESULT 14

STH0

somatotropin - horse

N:Alternate names: growth hormone

C:Species: Equus caballus (domestic horse)

C&gt;Date: 13-Jul-1981 #sequence\_revision 13-Jul-1981 #text\_change 23-Aug-1996

C:Accession: A01772; A91955; A01383; A90240; A01514

R:Zakari, M.M.; Poskus, E.; Langton, A.A.; Ferraira, P.; Santome, J.A.; Dellacha, J.M.; Pa

Int. J. Pept. Protein Res. 8, 435-444, 1976

A:Title: Primary structure of equine growth hormone.

A:Reference number: A91772; MUID:77005410; PMID:965151

A:Accession: A91772

A:Molecule type: protein

A:Residues: 1-190 &lt;ZAK&gt;

R:Zakari, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.

R:FEBS Lett. 34, 353-355, 1973

A:Title: The amino acid sequence of equine growth hormone.

A:Reference number: A91395; MUID:74020362; PMID:4747849

A:Accession: A91395

A:Molecule type: protein

A:Residues: 1-190 &lt;ZAK&gt;

R:Zakari, M.M.; Poskus, E.; Dellacha, J.M.; Paladini, A.C.; Santome, J.A.

R:FEBS Lett. 25, 77-82, 1972

A:Title: Amino acid sequences around the cysteine residues in equine growth hormone.

A:Reference number: A91383

A:Accession: A91383

A:Molecule type: protein

A:Residues: 42-69;157-190 &lt;ZAK&gt;

R:Oliver, J.; Hartree, A.S.

R:Biochem. J. 109, 19-24, 1968

A:Title: Amino acid sequences around the cysteine residues in horse growth hormone.

A:Reference number: A90240; MUID:68368390; PMID:4876100

A:Accession: A90240

A:Molecule type: protein

A:Residues: 176-190 &lt;OLI&gt;

C:Superfamily: prolactin

C:Keywords: hormone; pituitary

F:52-163,180-188/Disulfide bonds: #status experimental

Query Match

Best Local Similarity 38.5%; Score 306.5; DB 1; Length 190;

Matches 60; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

Db

2 FPTPLSLRFDNAMLRAHRLHQAFTYQEFEEAYIPKQKYSFLQNPQTSLSFSES IPT 61

1 FPMPLSSLSFANAVLRAQHQLADTYKEFERAYIPBQGRYS-IQNAQAFCSSETIPA 59

Qy

62 PSNRETOOKSNLELRISLLIQSWLEPVQ 93

60 PTGKEAQQRSDWELRLFSLLIQSWLGPVQ 91

Db

RESULT 15

STMS

somatotropin precursor - mouse

N:Alternate names: growth hormone

C:Species: Mus musculus (house mouse)

C&gt;Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 09-Jul-2004

C:Accession: B23911

R:Linzer, D.I.H.; Talamantes, F.

R:J. Biol. Chem. 260, 9574-9579, 1985

A:Title: Nucleotide sequence of mouse prolactin and growth hormone mRNAs and expression

A:Reference number: A92548; MUID:85261358; PMID:2991252

A:Accession: B23911

A:Molecule type: mRNA

A:Residues: 1-216 &lt;LIN&gt;

A:Cross-references: UNIPROT:P06880; GB:X02891; GB:X03232; NID:951067; PIDD:CAA26650.1;

C:Superfamily: prolactin

C:Keywords: anterior pituitary; growth factor; hormone

F:1-26/Domain: signal sequence #status predicted &lt;SIG&gt;

F:27-216/Product: somatotropin #status predicted &lt;STN&gt;

F:78-189,206-214/Disulfide bonds: #status predicted

Query Match

Best Local Similarity 38.2%; Score 304.5; DB 1; Length 216;

Matches 59; Conservative 14; Mismatches 17; Indels 1; Gaps 1;

Db

2 FPTPLSLRFDNAMLRAHRLHQAFTYQEFEEAYIPKQKYSFLQNPQTSLSFSES IPT 61

27 FPMPLSSLSFANAVLRAQHQLADTYKEFERAYIPBQGRYS-IQNAQAFCSSETIPA 85

Qy

62 PSNRETOOKSNLELRISLLIQSWLEPVQ 92

Sat Nov 6 18:59:33 2004

us-10-054-873-7.rpx

Page 7

DB 86 PRCRQAQRTDMLRPSLLIQSWLGPVQ 116

Search completed: November 2, 2004, 20:22:16  
JOB time : 27.6753 secs



**This Page Blank (uspto)**

Sat Nov 6 18:59:32 2004

us-10-054-873-7.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

CM protein - protein search, using sw model

Run on: November 2, 2004, 20:20:47 ; Search time 112.362 Seconds  
(without alignments)  
432.820 Million cell updates/sec

Title: US-10-054-873-7

Perfect score: 797  
Sequence: 1 MFPTPLSRFLPDNMLRAHR.....IVEQCTSTCSLYOLENYCN 150

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubppa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppa/US09A\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppa/US09B\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppa/US09D\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppa/US09E\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubppa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubppa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	797	100.0	150	US-10-054-873-7	Sequence 7, Appli
2	555.5	69.7	107	US-10-054-873-6	Sequence 6, Appli
3	470	59.0	92	US-10-054-873-2	Sequence 2, Appli
4	470	59.0	134	US-09-819-094-24	Sequence 24, Appli
5	470	59.0	134	US-10-714-067-24	Sequence 24, Appli
6	466	58.5	188	US-10-621-693-18	Sequence 18, Appli
7	466	58.5	192	US-09-819-094-23	Sequence 23, Appli
8	466	58.5	192	US-10-621-693-8	Sequence 8, Appli
9	466	58.5	192	US-10-621-693-78	Sequence 78, Appli
10	466	58.5	192	US-10-621-693-86	Sequence 86, Appli
11	466	58.5	192	US-10-714-067-23	Sequence 23, Appli
12	466	58.5	193	US-10-621-693-42	Sequence 42, Appli
13	466	58.5	206	US-10-621-693-72	Sequence 72, Appli

14	466	58.5	391	US-10-621-693-51	Sequence 51, Appli
15	466	58.5	574	US-10-621-693-32	Sequence 32, Appli
16	466	58.5	576	US-10-621-693-19	Sequence 39, Appli
17	466	58.5	599	US-10-621-693-53	Sequence 53, Appli
18	466	58.5	786	US-10-621-693-55	Sequence 55, Appli
19	466	58.5	810	US-10-621-693-76	Sequence 76, Appli
20	466	58.2	191	US-10-658-834A-875	Sequence 875, App
21	463	58.1	191	US-10-658-834A-866	Sequence 866, App
22	463	58.1	191	US-10-658-834A-876	Sequence 876, App
23	463	58.1	191	US-10-658-834A-887	Sequence 887, App
24	462	58.0	191	US-10-658-834A-867	Sequence 867, App
25	462	58.0	191	US-10-658-834A-881	Sequence 881, App
26	462	58.0	191	US-10-658-834A-888	Sequence 888, App
27	461	57.8	191	US-09-984-010-23	Sequence 23, Appli
28	461	57.8	191	US-10-153-207-1	Sequence 1, Appli
29	461	57.8	191	US-10-400-377-1	Sequence 1, Appli
30	461	57.8	191	US-10-400-708-1	Sequence 1, Appli
31	461	57.8	191	US-10-298-148-1	Sequence 1, Appli
32	461	57.8	191	US-10-646-798-2	Sequence 2, Appli
33	461	57.8	191	US-10-621-693-2	Sequence 2, Appli
34	461	57.8	191	US-10-621-693-21	Sequence 21, Appli
35	461	57.8	191	US-10-621-693-80	Sequence 80, Appli
36	461	57.8	191	US-10-621-693-82	Sequence 82, Appli
37	461	57.8	191	US-10-621-693-84	Sequence 84, Appli
38	461	57.8	191	US-10-718-340-1	Sequence 1, Appli
39	461	57.8	191	US-10-658-834A-868	Sequence 868, App
40	461	57.8	191	US-10-658-834A-869	Sequence 869, App
41	461	57.8	191	US-10-658-834A-870	Sequence 870, App
42	461	57.8	191	US-10-658-834A-871	Sequence 871, App
43	461	57.8	191	US-10-658-834A-883	Sequence 883, App
44	461	57.8	191	US-10-658-834A-884	Sequence 884, App
45	461	57.8	191	US-10-658-834A-885	Sequence 885, App

# ALIGNMENTS

RESULT 1  
US-10-054-873-7  
; Sequence 7, Application US/10054873  
; Publication No. US20020164712A1  
GENERAL INFORMATION:  
APPLICANT: Gan, Zhong Ru  
TITLE OF INVENTION: Chimeric Protein Containing an  
Intramolecular Chapterone-Like Sequence  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/054,873  
FILING DATE: 22-Jan-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/CN98/00052  
FILING DATE: 31-MAR-1998  
APPLICATION NUMBER: US 09/423,100  
FILING DATE: 11-DEC-2000  
ATTORNEY/AGENT INFORMATION:  
NAME: Mycroft, Frank J  
REGISTRATION NUMBER: 46,946  
REFERENCE/DOCKET NUMBER: 020167-0001300S  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:

LENGTH: 150 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 7  
US-10-054-873-7

Query Match 100.0%; Score 797; DB 13; Length 150;  
Best Local Similarity 100.0%; Pred. No. 2e-78;  
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSLRFDNMLRAHRLHQLAFDTYQFEFEAYIPKEOKYSFLQNPQTSLSFSSESIP 60  
DB 1 MFPTPLSLRFDNMLRAHRLHQLAFDTYQFEFEAYIPKEOKYSFLQNPQTSLSFSSESIP 60  
QY 61 TPSNEETQOKSNLELRISILLIQSWLEPVLGTGPRFVNHLCGSHLVEALYVCGER 120  
DB 61 TPSNEETQOKSNLELRISILLIQSWLEPVLGTGPRFVNHLCGSHLVEALYVCGER 120  
QY 121 GFFYTPKTRGIVEOCCTISCSLYOLENYCN 150  
DB 121 GFFYTPKTRGIVEOCCTISCSLYOLENYCN 150

## RESULT 2

US-10-054-873-6  
Sequence 6, Application US/10054873  
Publication No. US20020164712A1

## GENERAL INFORMATION:

APPLICANT: Gan, Zhong Ru  
TITLE OF INVENTION: Chimeric Protein Containing an  
Intramolecular Chaperone-Like Sequence  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/054,873

FILING DATE: 22-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: NO PCT/CN98/00052

FILING DATE: 31-MAR-1998

APPLICATION NUMBER: US 09/423,100

FILING DATE: 11-DEC-2000

ATTORNEY/AGENT INFORMATION:

NAME: Mycroft, Frank J

REGISTRATION NUMBER: 46,946

REFERENCE/DOCKET NUMBER: 020167-000130US

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 107 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-10-054-873-6

Query Match 69.7%; Score 555.5; DB 13; Length 107;  
Best Local Similarity 71.3%; Pred. No. 2.3e-52;  
Matches 107; Conservative 0; Mismatches 0; Indels 43; Gaps 1;

QY 1 MFPTPLSLRFDNMLRAHRLHQLAFDTYQFEFEAYIPKEOKYSFLQNPQTSLSFSSESIP 60  
DB 1 MFPTPLSLRFDNMLRAHRLHQLAFDTYQFEFEAYIPKEOKYSFLQNP----- 49  
QY 61 TPSNEETQOKSNLELRISILLIQSWLEPVLGTGPRFVNHLCGSHLVEALYVCGER 120  
DB 50 -----LGTGPRFVNHLCGSHLVEALYVCGER 77

QY 121 GFFYTPKTRGIVEOCCTISCSLYOLENYCN 150  
DB 78 GFFYTPKTRGIVEOCCTISCSLYOLENYCN 107

## RESULT 3

US-10-054-873-2  
Sequence 2, Application US/10054873  
Publication No. US20020164712A1

## GENERAL INFORMATION:

APPLICANT: Gan, Zhong Ru  
TITLE OF INVENTION: Chimeric Protein Containing an  
Intramolecular Chaperone-Like Sequence  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:

ADDRESS: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/054,873

FILING DATE: 22-Jan-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: NO PCT/CN98/00052

FILING DATE: 31-MAR-1998

APPLICATION NUMBER: US 09/423,100

FILING DATE: 11-DEC-2000

ATTORNEY/AGENT INFORMATION:

NAME: Mycroft, Frank J

REGISTRATION NUMBER: 46,946

REFERENCE/DOCKET NUMBER: 020167-000130US

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 92 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-054-873-2

Query Match 59.0%; Score 470; DB 13; Length 92;  
Best Local Similarity 100.0%; Pred. No. 3.8e-43;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPTPLSLRFDNMLRAHRLHQLAFDTYQFEFEAYIPKEOKYSFLQNPQTSLSFSSESIP 60  
DB 1 MFPTPLSLRFDNMLRAHRLHQLAFDTYQFEFEAYIPKEOKYSFLQNPQTSLSFSSESIP 60  
QY 61 TPSNEETQOKSNLELRISILLIQSWLEPVLGTGPRFVNHLCGSHLVEALYVCGER 120  
DB 61 TPSNEETQOKSNLELRISILLIQSWLEPVLGTGPRFVNHLCGSHLVEALYVCGER 120

## RESULT 4

US-09-819-094-24  
Sequence 24, Application US/09819094

Publication No. US20030186382A1  
GENERAL INFORMATION:  
APPLICANT: Weiner, Richard I.  
APPLICANT: Martini, Joseph A.  
APPLICANT: Struman, Ingrid  
APPLICANT: Taylor, Robert  
APPLICANT: Bentzen, Frauke  
TITLE OF INVENTION: No. US20030186382A1 Antiangiogenic Peptide Agents and Their  
TITLE OF INVENTION: Therapeutic and Diagnostic Use  
FILE REFERENCE: USCF-018/02US  
CURRENT APPLICATION NUMBER: US/09/819,094  
PRIORITY FILING DATE: 2001-03-27  
PRIORITY FILING DATE: 1998-05-12  
PRIORITY FILING DATE: 1998-05-12  
PRIORITY FILING DATE: 1997-05-12  
NUMBER OF SEQ ID NOS: 34  
SEQ ID NO 24  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-819-094-24

Query Match 59.0%; Score 470; DB 10; Length 134;  
Best Local Similarity 100.0%; Pred. No. 6.1e-43;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTTIPSLRFLPDNMLRAHRLHQLAFDPTYOEFEEAYIPKQKYSFLQNPOTSLSFSSSIP 60  
DB 1 MPTTIPSLRFLPDNMLRAHRLHQLAFDPTYOEFEEAYIPKQKYSFLQNPOTSLSFSSSIP 60  
QY 61 TPSNREETOQKSNLELRISILLIOSWLEPVQ 92  
DB 61 TPSNREETOQKSNLELRISILLIOSWLEPVQ 92

RESULT 5  
US-10-714-067-24  
Sequence 24, Application US/10714067  
Publication No. US20040077054A1  
GENERAL INFORMATION:  
APPLICANT: Weiner, Richard I.  
APPLICANT: Martini, Joseph A.  
APPLICANT: Struman, Ingrid  
APPLICANT: Taylor, Robert  
APPLICANT: Bentzen, Frauke  
TITLE OF INVENTION: Novel Antiangiogenic Peptide Agents and Their  
TITLE OF INVENTION: Therapeutic and Diagnostic Use  
FILE REFERENCE: USCF-018/02US  
CURRENT APPLICATION NUMBER: US/10/714,067  
PRIORITY FILING DATE: 2003-11-14  
PRIORITY FILING DATE: 2001-03-27  
PRIORITY FILING DATE: 1998-05-12  
PRIORITY FILING DATE: 1997-05-12  
NUMBER OF SEQ ID NOS: 34  
SEQ ID NO 24  
LENGTH: 134  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-714-067-24

Query Match 59.0%; Score 470; DB 15; Length 134;  
Best Local Similarity 100.0%; Pred. No. 6.1e-43;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPTTIPSLRFLPDNMLRAHRLHQLAFDPTYOEFEEAYIPKQKYSFLQNPOTSLSFSSSIP 60  
DB 1 MPTTIPSLRFLPDNMLRAHRLHQLAFDPTYOEFEEAYIPKQKYSFLQNPOTSLSFSSSIP 60  
QY 61 TPSNREETOQKSNLELRISILLIOSWLEPVQ 92

DB 61 TPSNREETOQKSNLELRISILLIOSWLEPVQ 92

RESULT 6  
US-10-621-693-18  
Sequence 18, Application US/10621693  
Publication No. US20040059053A1  
GENERAL INFORMATION:  
APPLICANT: Bussell, Stuart  
APPLICANT: Genetec Biopharmaceuticals, Inc.  
TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQU  
TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS  
FILE REFERENCE: GNT-00101.P.1-US  
CURRENT APPLICATION NUMBER: US/10/621,693  
PRIORITY FILING DATE: 2003-07-16  
PRIORITY FILING DATE: 2002-07-16  
PRIORITY FILING DATE: 2002-07-16  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 18  
LENGTH: 188  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: synthetic sequence  
US-10-621-693-18

Query Match 58.5%; Score 466; DB 15; Length 188;  
Best Local Similarity 70.5%; Pred. No. 2.6e-42;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MPTTIPSLRFLPDNMLRAHRLHQLAFDPTYOEFEEAYIPKQKYSFLQNPOTSLSFSSSIP 60  
DB 1 MPTTIPSLRFLPDNMLRAHRLHQLAFDPTYOEFEEAYIPKQKYSFLQNPOTSLSFSSSIP 60  
QY 61 TPSNREETOQKSNLELRISILLIOSWLEPVQ 92  
DB 61 TPSNREETOQKSNLELRISILLIOSWLEPVQ 92  
QY 111 BALYVCG-ERGFYTPKTRGIVEQ 134  
DB 120 EGIOTMGRLEBDG--SPRTGQIFKQ 142

RESULT 7  
US-09-819-094-23  
Sequence 23, Application US/09819094  
Publication No. US20030186382A1  
GENERAL INFORMATION:  
APPLICANT: Weiner, Richard I.  
APPLICANT: Martini, Joseph A.  
APPLICANT: Struman, Ingrid  
APPLICANT: Taylor, Robert  
APPLICANT: Bentzen, Frauke  
TITLE OF INVENTION: No. US20030186382A1 Antiangiogenic Peptide Agents and Their  
TITLE OF INVENTION: Therapeutic and Diagnostic Use  
FILE REFERENCE: USCF-018/02US  
CURRENT APPLICATION NUMBER: US/09/819,094  
PRIORITY FILING DATE: 2001-03-27  
PRIORITY FILING DATE: 1998-05-12  
PRIORITY FILING DATE: 1997-05-12  
NUMBER OF SEQ ID NOS: 34  
SEQ ID NO 23  
LENGTH: 192  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-819-094-23

Query Match 58.5%; Score 466; DB 10; Length 192;  
Best Local Similarity 70.5%; Pred. No. 2.7e-42;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

Db 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

QY 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

Db 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

QY 111 EALYLVCG--ERGFYTPKTRGIVEQ 134

Db 120 EGIQTLMGRLBDG---SPRTGQIFKQ 142

#### RESULT 8

US-10-621-693-8

Sequence 8, Application US/10621693

Publication No. US20040059093A1

GENERAL INFORMATION:

APPLICANT: Gentide Biopharmaceuticals, Inc.

APPLICANT: Bussell, Stuart

TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENCES

FILE REFERENCE: GNT-00101.P.1-US

CURRENT FILING DATE: 2003-07-16

PRIOR APPLICATION NUMBER: US 60/396,466

PRIOR FILING DATE: 2002-07-16

NUMBER OF SEQ ID NOS: 86

SOFTWARE: Patent in version 3.0

SEQ ID NO 8

LENGTH: 192

TYPE: PRT

ORGANISM: Artificial

OTHER INFORMATION: synthetic sequence

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: (1)..()

US-10-621-693-8

Query Match

Best Local Similarity 58.5%; Score 466; DB 15; Length 192;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

Db 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

QY 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

Db 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

QY 111 EALYLVCG--ERGFYTPKTRGIVEQ 134

Db 120 EGIQTLMGRLBDG---SPRTGQIFKQ 142

#### RESULT 9

US-10-621-693-78

Sequence 78, Application US/10621693

Publication No. US20040059093A1

GENERAL INFORMATION:

APPLICANT: Gentide Biopharmaceuticals, Inc.

APPLICANT: Bussell, Stuart

TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENCES

FILE REFERENCE: GNT-00101.P.1-US

CURRENT FILING DATE: 2003-07-16

PRIOR APPLICATION NUMBER: US 60/396,466

PRIOR FILING DATE: 2002-07-16

NUMBER OF SEQ ID NOS: 86

SOFTWARE: Patent in version 3.0

SEQ ID NO 78

LENGTH: 192

TYPE: PRT

ORGANISM: Artificial

OTHER INFORMATION: synthetic sequence

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: (1)..()

US-10-621-693-78

Query Match

Best Local Similarity 58.5%; Score 466; DB 15; Length 192;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

Db 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

QY 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

Db 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

QY 111 EALYLVCG--ERGFYTPKTRGIVEQ 134

Db 120 EGIQTLMGRLBDG---SPRTGQIFKQ 142

#### RESULT 10

US-10-621-693-86

Sequence 86, Application US/10621693

Publication No. US20040059093A1

GENERAL INFORMATION:

APPLICANT: Gentide Biopharmaceuticals, Inc.

APPLICANT: Bussell, Stuart

TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENCES

FILE REFERENCE: GNT-00101.P.1-US

CURRENT FILING DATE: 2003-07-16

PRIOR APPLICATION NUMBER: US 60/396,466

PRIOR FILING DATE: 2002-07-16

NUMBER OF SEQ ID NOS: 86

SOFTWARE: Patent in version 3.0

SEQ ID NO 86

LENGTH: 192

TYPE: PRT

ORGANISM: Artificial

OTHER INFORMATION: synthetic sequence

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (2)..(192)

OTHER INFORMATION: sequence is repeated N+2 times, where N is a positive whole number

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: (1)..()

US-10-621-693-86

Query Match

Best Local Similarity 58.5%; Score 466; DB 15; Length 192;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

Db 1 MFPTPLSLFNDNMLRAHRLHQLAFDTYQEFEEBAYIPKQKYSFLONPOTSLSFSESIP 60

QY 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

Db 61 TFSNREETOQKSNLELIRISILLIOSWLEPVQLGTGRPFVNOHLCGS-----HLV 110

QY 111 EALYLVCG--ERGFYTPKTRGIVEQ 134

Db 120 EGIQTLMGRLBDG---SPRTGQIFKQ 142

```
RESULT 11
US-10-714-067-23
; Sequence 23, Application US/10714067
; Publication No. US2004007054A1
; GENERAL INFORMATION:
; APPLICANT: Weiner, Richard I.
; APPLICANT: Marital, Joseph A.
; APPLICANT: Struman, Ingrid
; APPLICANT: Taylor, Robert
; APPLICANT: Benizien, Frauke
; TITLE OF INVENTION: Novel Antiangiogenic Peptide Agents and Their
; FILE REFERENCE: UCSF-018/0205
; CURRENT APPLICATION NUMBER: US/10/714,067
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US/09/819,094
; PRIOR FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 09/076,675
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: 60/046,394
; PRIOR FILING DATE: 1997-05-12
; NUMBER OF SEQ. ID NOS: 34
; SEQ ID NO 23
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-714-067-23

Query Match
Best Local Similarity 58.5%; Score 466; DB 15; Length 192;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFTPIPSRLFDNMLRAHRLHQLAPDTYQEFEEAYIPKQKXSFLONPQTSLSFSESIP 60
DB 1 MFTPIPSRLFDNMLRAHRLHQLAPDTYQEFEEAYIPKQKXSFLONPQTSLSFSESIP 60
QY 61 TPSNREETOQKSNLELRISLLIQSWLEPVQLGTGPRFVFNCHLGS-----HLV 110
DB 61 TPSNREETOQKSNLELRISLLIQSWLEPVQLGTGPRFVFNCHLGS-----HLV 110
QY 111 EALYIVCG--ERGFYTPKTRGIVEQ 134
DB 120 EGIQTLMGRLDQ--SPRTGQIFKQ 142

RESULT 12
US-10-621-693-42
; Sequence 42, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Genetide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENC
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; PRIOR FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
US-10-621-693-42

Query Match
Best Local Similarity 58.5%; Score 466; DB 15; Length 193;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFTPIPSRLFDNMLRAHRLHQLAPDTYQEFEEAYIPKQKXSFLONPQTSLSFSESIP 60
DB 1 MFTPIPSRLFDNMLRAHRLHQLAPDTYQEFEEAYIPKQKXSFLONPQTSLSFSESIP 60
QY 61 TPSNREETOQKSNLELRISLLIQSWLEPVQLGTGPRFVFNCHLGS-----HLV 110
DB 61 TPSNREETOQKSNLELRISLLIQSWLEPVQLGTGPRFVFNCHLGS-----HLV 110
QY 111 EALYIVCG--ERGFYTPKTRGIVEQ 134
DB 120 EGIQTLMGRLDQ--SPRTGQIFKQ 142

RESULT 13
US-10-621-693-72
; Sequence 72, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Genetide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUE
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; PRIOR FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: synthetic sequence
US-10-621-693-72

Query Match
Best Local Similarity 70.5%; Score 466; DB 15; Length 206;
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFTPIPSRLFDNMLRAHRLHQLAPDTYQEFEEAYIPKQKXSFLONPQTSLSFSESIP 60
DB 1 MFTPIPSRLFDNMLRAHRLHQLAPDTYQEFEEAYIPKQKXSFLONPQTSLSFSESIP 60
QY 61 TPSNREETOQKSNLELRISLLIQSWLEPVQLGTGPRFVFNCHLGS-----HLV 110
DB 61 TPSNREETOQKSNLELRISLLIQSWLEPVQLGTGPRFVFNCHLGS-----HLV 110
QY 111 EALYIVCG--ERGFYTPKTRGIVEQ 134
DB 120 EGIQTLMGRLDQ--SPRTGQIFKQ 142

RESULT 14
US-10-621-693-51
; Sequence 51, Application US/10621693
; Publication No. US20040059093A1
; GENERAL INFORMATION:
; APPLICANT: Genetide Biopharmaceuticals, Inc.
; APPLICANT: Bussell, Stuart
; TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENC
; FILE REFERENCE: GNT-00101.P.1-US
; CURRENT APPLICATION NUMBER: US/10/621,693
; PRIOR FILING DATE: 2003-07-16
; PRIOR APPLICATION NUMBER: US 60/396,466
; PRIOR FILING DATE: 2002-07-16
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51
```

Search completed: November 2, 2004, 20:59:23  
Job time : 112.362 secs

LENGTH: 391  
TYPE: PRT  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: synthetic sequence  
FEATURE:  
NAME/KEY: mac\_peptide  
LOCATION: (1)..()  
US-10-621-693-51

Query Match 58.5%; Score 466; DB 15; Length 391;  
Best Local Similarity 70.5%; Pred. No. 6,7e-42;  
Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSRFLFDNMLRAHRLHQLAFTYQFEFEAYIPKEOKYSFLONPQTSLSFSSESIP 60  
DB 1 MFPTPLSRFLFDNMLRAHRLHQLAFTYQFEFEAYIPKEOKYSFLONPQTSLSFSSESIP 60  
QY 61 TPSNREETOQKSNLELRISILLIQSWLEBPVQGTGPRFVNOHLGCS-----HLV 110  
DB 61 TPSNREETOQKSNLELRISILLIQSWLEBPVQGTGPRFVNOHLGCS-----HLV 110  
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
QY 120 EGIQTLMGRLSDG---SPRTGQIFKQ 142  
DB 120 EGIQTLMGRLSDG---SPRTGQIFKQ 142

## RESULT 15

US-10-621-693-32  
Sequence 32, Application US/10621693  
Publication No. US20040059093A1

GENERAL INFORMATION:

APPLICANT: Genetide Biopharmaceuticals, Inc.

APPLICANT: Bussell, Stuart

TITLE OF INVENTION: METHODS TO CONSTRUCT MULTIMERIC DNA AND POLYMERIC PROTEIN SEQUENC

TITLE OF INVENTION: DIRECT FUSIONS OR WITH LINKERS

FILE REFERENCE: GNT-00101.P.1-US

CURRENT APPLICATION NUMBER: US/10/621,693

PRIOR FILING DATE: 2003-07-16

PRIOR APPLICATION NUMBER: US 60/396,466

NUMBER OF SEQ ID NOS: 86

SOFTWARE: PatentIn version 3.0

SEQ ID NO 32

LENGTH: 574

TYPE: PRT

ORGANISM: Artificial

FEATURE:

OTHER INFORMATION: synthetic sequence

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (379)..(569)

OTHER INFORMATION: sequence is repeated N-1 times, where N is a positive whole numbe

FEATURE:

NAME/KEY: mac\_peptide

LOCATION: (1)..()

US-10-621-693-32

Query Match 58.5%; Score 466; DB 15; Length 574;

Best Local Similarity 70.5%; Pred. No. 1.1e-41;

Matches 103; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 1 MFPTPLSRFLFDNMLRAHRLHQLAFTYQFEFEAYIPKEOKYSFLONPQTSLSFSSESIP 60  
DB 1 MFPTPLSRFLFDNMLRAHRLHQLAFTYQFEFEAYIPKEOKYSFLONPQTSLSFSSESIP 60  
QY 61 TPSNREETOQKSNLELRISILLIQSWLEBPVQGTGPRFVNOHLGCS-----HLV 110  
DB 61 TPSNREETOQKSNLELRISILLIQSWLEBPVQGTGPRFVNOHLGCS-----HLV 110  
QY 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
DB 111 EALYVCG--ERGFYTPKTRGIVEQ 134  
QY 120 EGIQTLMGRLSDG---SPRTGQIFKQ 142  
DB 120 EGIQTLMGRLSDG---SPRTGQIFKQ 142



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: November 2, 2004, 19:48:36 ; Search time 145.295 Seconds  
(without alignments)  
594,006 Million cell updates/sec

Title: US-10-054-873-7

Perfect score: 797

Sequence: 1 MFPTPLSRLEFDMNMAHR.....IVROCTSTICSLYQLENYCN 150

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	461	57.8	217	1	SOMA_HUMAN
2	461	57.8	217	1	SOMA_PANTR
3	460	57.7	217	1	SOMA_MACMU
4	460	57.7	217	2	CELYFO
5	460	57.7	217	2	AA11509
6	457	57.3	217	2	CELYFO
7	457	57.3	217	2	AA11508
8	457	57.3	217	1	SOMA_SALTB
9	456	54.7	217	1	QWNEO
10	434.5	54.5	217	1	SOMA_PANTR
11	432	54.2	217	1	SOMA_CALVA
12	427.5	53.6	217	1	CEH54
13	427.5	53.6	245	2	014644
14	426.5	53.5	217	1	SOMA2_HUMAN
15	421.5	52.9	217	2	CEH52
16	407.5	51.1	217	2	QW369
17	399	50.1	184	2	QW679
18	399	50.1	184	1	SOMA2_MACMU
19	397	49.8	217	2	QW6679
20	396	49.7	212	2	QW368
21	396	49.7	217	2	QW367
22	385	48.3	217	2	QW6678
23	381	47.8	217	1	CSH_HUMAN
24	381	47.8	217	2	QW671
25	381	47.8	217	2	AAH57768
26	381	47.8	217	2	AAH57768
27	381	47.8	217	2	AAH57768
28	370	46.4	217	2	QW6670
29	365.5	45.9	202	2	AAH62475
30	348	43.7	217	2	QW6670
31	341	42.8	202	2	014643

32	322.5	40.5	217	2	QW67174	QW67174 callitrix
33	310.5	39.0	216	1	SOMA_MESAU	P37866 mesocricetu
34	307.5	38.6	190	1	SOMA_BALBO	P33092 balaeopreter
35	306.5	38.5	216	1	SOMA_HORSE	P01245 equus caball
36	306.5	38.5	216	2	070615	O70615 spalax leuc
37	306.5	38.5	217	1	SOMA_GALSE	Q99841 galago sene
38	306.5	38.5	217	1	SOMA_NICPY	Q99842 nycticebus
39	304.5	38.2	216	1	SOMA_MOUSE	P06880 mus musculu
40	304.5	38.2	216	2	BAB31932	Bab31932 mus muscu
41	304.5	38.2	216	2	BAB31933	Bab31933 mus muscu
42	304.5	38.2	216	2	BAB31935	Bab31935 mus muscu
43	304.5	38.2	216	2	BAB31937	Bab31937 mus muscu
44	304.5	38.2	216	2	BAC27096	Bac27096 mus muscu
45	302.5	38.0	216	1	SOMA_RABIT	P46407 oryctolagus

## ALIGNMENTS

RESULT 1  
SOMA\_HUMAN STANDARD: PRT: 217 AA.  
ID P01241; 014405; Q16631; Q9HEZ1; Q9UMW7; Q9UNL5;  
AC 21-TU-1866 (Rel. 01, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE Somatotropin precursor (growth hormone) (GH) (GH-N) (pituitary growth hormone) (growth hormone 1).  
GN Name=GH1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OK NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=80034477; PubMed=386281;  
RA Roskam W., Rougeon F.;  
RT "Molecular cloning and nucleotide sequence of the human growth hormone structural gene."  
RL Nucleic Acids Res. 7:305-320(1979).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RX MEDLINE=79203293; PubMed=377496;  
RA Martini J.A., Hallett R.A., Baxter J.D., Goodman H.M.;  
RT "Human growth hormone: complementary DNA cloning and expression in bacteria."  
RL Science 205:602-607(1979).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND POSSIBLE ALTERNATIVE SPLICING.  
RX MEDLINE=82014939; PubMed=6269091;  
RA Denoto F.M., Moore D.D., Goodman H.M.;  
RT "Human growth hormone DNA sequence and mRNA structure: possible alternative splicing."  
RL Nucleic Acids Res. 9:3719-3730(1981).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8182010; PubMed=7169009;  
RA Seedburg P.H.;  
RT "The human growth hormone gene family: nucleotide sequences show recent divergence and predict a new polypeptide hormone."  
RL DNA 1:239-249(1982).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89307277; PubMed=2744760;  
RA Chen B.Y., Liao Y.C., Smith D.H., Barrera-Saldana H.A., Gellinas R.E., Seedburg P.H.;  
RT "The human growth hormone locus: nucleotide sequence, biology, and evolution."  
RL Genomics 4:479-497(1989).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORM 3).  
RX TISST5=Plutitary;  
GU J., Huang Q.-H., Li N., Xu S.-H., Han Z.-G., Fu G., Chen Z.;

- RT "A novel gene expressed in human pituitary.";  
RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [17]  
RC SEQUENCE FROM N.A. (ISOFORM 4).  
RA MEDLINE=20402571; PubMed=10931946;  
RA Hu R.-W., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,  
RA Yu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,  
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,  
RA Xu S.-H., Gu J.-X., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,  
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;  
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal  
RT axis and full-length cDNA cloning.";  
RT Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).  
RN [18]  
RP SEQUENCE OF 1-26 FROM N.A.  
RC MEDLINE=86137393; PubMed=3912261;  
RA Gray G.L., Baldridge J.S., McKewen K.S., Heyneker H.L., Chang C.N.;  
RT "Periplasmic production of correctly processed human growth hormone in  
RT *Escherichia coli*: natural and bacterial signal sequences are  
RT interchangeable.";  
RT Gene 39:247-254(1995).  
RN [19]  
RP SEQUENCE OF 27-217.  
RC MEDLINE=69289202; PubMed=5810834;  
RA Li C.H., Dixon J.S., Liu W.-K.;  
RT "Human pituitary growth hormone. XIX. The primary structure of the  
RT hormone.";  
RT Arch. Biochem. Biophys. 133:70-91(1969).  
RN [110]  
RP SEQUENCE OF 27-217, AND REVISIONS.  
RC MEDLINE=72143935; PubMed=5144027;  
RA Li C.H., Dixon J.S.;  
RT "Human pituitary growth hormone. 32. The primary structure of the  
RT hormone: revision.";  
RT Arch. Biochem. Biophys. 146:233-236(1971).  
RN [111]  
RP REVISION.  
RC MEDLINE=73092028; PubMed=4675454;  
RA Bewley T.A., Dixon J.S., Li C.H.;  
RT "Sequence comparison of human pituitary growth hormone, human  
RT chorionic somatomammotropin, and ovine pituitary growth and lactogenic  
RT hormones.";  
RT Int. J. Pept. Protein Res. 4:281-287(1972).  
RN [112]  
RP SEQUENCE OF 27-61 AND 102-124.  
RC MEDLINE=71139765; PubMed=5279046;  
RA Niall H.D.;  
RT "Revised primary structure for human growth hormone.";  
RT Nature New Biol. 230:90-91(1971).  
RN [113]  
RP REVISIONS TO 119-120 AND 157-159.  
RC MEDLINE=71153968; PubMed=5279528;  
RA Niall H.D., Hogan M.L., Sauer R., Rosenblum I.Y., Greenwood F.C.;  
RT "Sequences of pituitary and placental lactogenic and growth hormones:  
RT evolution from a primordial peptide by gene reduplication.";  
RT Proc. Natl. Acad. Sci. U.S.A. 68:866-869(1971).  
RN [114]  
RP REVISION.  
RA Niall H.D.;  
RT "The chemistry of the human lactogenic hormones.";  
RT (In) Griffiths K. (eds.);  
RT Prolactin and carcinogenesis. Proc. fourth tenous workshop prolactin,  
RT pp.13-20, Alpha Omega Alpha Press, Cardiff (1972).  
RN [115]  
RP SEQUENCE OF 27-79 (ISOFORM 2).  
RC MEDLINE=8117361; PubMed=7462247;  
RA Chapman G.E., Rogers K.M., Brittain T., Bradshaw R.A., Bates O.J.,  
RA Turner C., Cary P.D., Crane-Robinson C.;  
RT "The 20,000 molecular weight variant of human growth hormone.  
RT Preparation and some physical and chemical properties.";  
RT J. Biol. Chem. 256:2395-2401(1981).  
RN [116]  
RP SEQUENCE OF 46-80 (ISOFORM 2).  
RX MEDLINE=80130196; PubMed=7356479;  
RA Lewis U.J., Bonewald L.F., Lewis L.J.;  
RT "The 20,000-dalton variant of human growth hormone: location of the  
RT amino acid deletions.";  
RT Biochem. Biophys. Res. Commun. 92:511-516(1980).  
RN [117]  
RP DEAMINATION OF GUN-163 AND ASN-178.  
RX MEDLINE=82052997; PubMed=7028740;  
RA Lewis U.J., Singh R.N., Bonewald L.F., Seavey B.K.;  
RT "Altered proteolytic cleavage of human growth hormone as a result of  
RT deamination.";  
RT J. Biol. Chem. 256:11645-11650(1981).  
RN [118]  
RP PHOSPHORYLATION SITES SER-132 AND SER-176.  
RC TISSUE=Pituitary;  
RC PubMed=14997482; DOI=10.1002/pmic.200300584;  
RA Giorgianni F., Beranova-Giorgianni S., Desiderio D.M.;  
RT "Identification and characterization of phosphorylated proteins in the  
RT human pituitary.";  
RT Proteomics 4:587-598(2004).  
RN [119]  
RP REVIEW.  
RC MEDLINE=99321812; PubMed=10393484;  
RA Baumann G.;  
RT "Growth hormone heterogeneity in human pituitary and plasma.";  
RT Horm. Res. 51 Suppl. 1:2-6(1999).  
RN [120]  
RP 3D-STRUCTURE MODELING.  
RX MEDLINE=88190073; PubMed=3447173;  
RA Cohen P.E., Kuntz I.D.;  
RT "Prediction of the three-dimensional structure of human growth  
RT hormone.";  
RT Proteins 2:162-166(1987).  
RN [121]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=92196577; PubMed=1549776;  
RA de Vos A.M., Ullrich M., Kossiakoff A.A.;  
RT "Human growth hormone and extracellular domain of its receptor:  
RT crystal structure of the complex.";  
RT Science 255:306-312(1992).  
RN [122]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE=95075462; PubMed=7984244;  
RA Somers W., Ullrich M., de Vos A.M., Kossiakoff A.A.;  
RT "The X-ray structure of a growth hormone-prolactin receptor complex.";  
RT Nature 372:478-481(1994).  
RN [123]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RA Chantrelat L., Chirgatz N.Y., Jones N., Korber F., Navaza J.,  
RA Pavlovsk A.G., Wlodawer A.;  
RT "The crystal-structure of wild-type growth-hormone at 2.5-A  
RT resolution.";  
RT Protein Pept. Lett. 2:333-340(1995).  
RN [124]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).  
RX MEDLINE=97113023; PubMed=8943276;  
RA Sundstroem M., Lundqvist T., Roedin J., Giebel L.B., Milligan D.,  
RA Norstedt G.;  
RT "Crystal structure of an antagonist mutant of human growth hormone,  
RT G120R, in complex with its receptor at 2.9-A resolution.";  
RT J. Biol. Chem. 271:32197-32203(1996).  
RN [125]  
RP VARIANT KOWARSKI SYNDROME CVS-103.  
RX MEDLINE=96150232; PubMed=8552145;  
RA Takahashi Y., Kajii H., Okimura Y., Goji K., Abe H., Chihara K.;  
RT "Short stature caused by a mutant growth hormone.";  
RT N. Engl. J. Med. 334:432-436(1996).  
RN [126]  
RP ERRATUM.  
RA Takahashi Y., Kajii H., Okimura Y., Goji K., Abe H., Chihara K.;  
RT N. Engl. J. Med. 334:1207-1207(1996).  
RN [127]

RP VARIANT KOMARSKI SYNDROME GLX-138.  
RX MEDLINE=97426478; PubMed=9276733;

Query Match 57.8%; Score 461; DB 1; Length 217;  
Best Local Similarity 70.3%; Pred. No. 1.7e-36;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTPLSLRFDNMLRAHRLHQLAFDTQCEFEBAVYIPKQKXSFLLQNPQTSLSFSISPT 61  
DB 27 PPTPLSLRFDNMLRAHRLHQLAFDTQCEFEBAVYIPKQKXSFLLQNPQTSLSFSISPT 86  
QY 62 PSNREETOQKSNLELIRISLLILQSWLEPVQGTGPFVQNHCGS-----HYVE 111  
DB 87 PSNREETOQKSNLELIRISLLILQSWLEPVQGTGPFVQNHCGS-----HYVE 145  
QY 112 ALYVCG--ERGFYTPKRGIVEQ 134  
DB 146 GIQTMGRLEDC--SPRTGQIFKQ 167

## RESULT 2

SOMA\_PANTR STANDARD; PRT; 217 AA.  
AC P58756;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).  
GN Name=GH1;  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
OX NCBI\_Taxid=9558;  
RN [1]  
RA SEQUENCE FROM N.A.  
RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;  
RT "Independent duplication of the growth hormone gene in three anthropoid lineages";  
RT Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the somatotropin/prolactin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; AF374232; AAL72264.1; -  
DR HSSP; P01241; IHWG.  
DR InterPro; IPR009079; 4\_helix\_cytokine.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00103; Hormone\_1; 1.  
DR PRINTS; PR00836; SOMATOTROPIN.  
DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
KM Hormone; Pituitary; Signal.  
FT SIGNAL 1  
FT CHAIN 27  
FT DISULFID 79  
FT DISULFID 191  
FT DISULFID 208  
FT DISULFID 215  
SQ SEQUENCE 217 AA; 24643 MW; FEA235ED0518674 CRC64;  
Query Match 57.8%; Score 461; DB 1; Length 217;

Best Local Similarity 70.3%; Pred. No. 1.7e-36;  
Matches 102; Conservative 7; Mismatches 20; Indels 16; Gaps 4;

QY 2 PPTPLSLRFDNMLRAHRLHQLAFDTQCEFEBAVYIPKQKXSFLLQNPQTSLSFSISPT 61  
DB 27 PPTPLSLRFDNMLRAHRLHQLAFDTQCEFEBAVYIPKQKXSFLLQNPQTSLSFSISPT 86  
QY 62 PSNREETOQKSNLELIRISLLILQSWLEPVQGTGPFVQNHCGS-----HYVE 111  
DB 87 PSNREETOQKSNLELIRISLLILQSWLEPVQGTGPFVQNHCGS-----HYVE 145  
QY 112 ALYVCG--ERGFYTPKRGIVEQ 134  
DB 146 GIQTMGRLEDC--SPRTGQIFKQ 167

## RESULT 3

SOMA\_MACMU STANDARD; PRT; 217 AA.  
AC P33093;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Somatotropin precursor (Growth hormone) (GH) (GH-N) (Pituitary growth hormone) (Growth hormone 1).  
GN Name=GH1;  
OS Macaca mulatta (Rhesus macaque).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;  
OC Cercopitheidae; Macaca.  
OX NCBI\_Taxid=9544;  
RN [1]  
RA SEQUENCE FROM N.A.  
RA MEDLINE=9408724; PubMed=8404617;  
RA Goios T.G., Durning M., Fisher J.M., Fowler P.D.;  
RT "Cloning of four growth hormone/chorionic somatomotropin-related complementary deoxyribonucleic acids differentially expressed during pregnancy in the rhesus monkey placenta";  
RT Endocrinology 133:1744-1752(1993).  
RN [2]  
RP SEQUENCE OF 27-217.  
RX MEDLINE=86129460; PubMed=3080959;  
RX Li C.H., Chung D., Lam H.W., Stein S.;  
RA "The primary structure of monkey pituitary growth hormone";  
RL Arch. Biochem. Biophys. 245:287-291(1986).  
CC -!- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- SIMILARITY: Belongs to the somatotropin/prolactin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; L16556; AAA18842.1; -  
DR PIR; I67410.  
DR HSSP; P01241; IAXI.  
DR InterPro; IPR009079; 4\_helix\_cytokine.  
DR InterPro; IPR001400; Somatotropin.  
DR Pfam; PF00103; Hormone\_1; 1.  
DR PRINTS; PR00836; SOMATOTROPIN.  
DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
KM Direct protein sequencing; Hormone; Pituitary; Signal.  
FT SIGNAL 1  
FT CHAIN 26

FT CHAIN 27 217 Somatotropin.  
 FT DISULFID 79 191 By similarity.  
 FT DISULFID 208 215 E -> O (in Ref. 2).  
 FT CONFLICT 100 100 N -> D (in Ref. 2).  
 FT CONFLICT 179 179  
 SQ SEQUENCE 217 AA; 24913 MW; 2C5180341EEC46D0 CRC64;

Query Match 57.7%; Score 460; DB 2; Length 217;  
 Best Local Similarity 98.9%; Pred. No. 2,1e-36;  
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 61  
 DB 27 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 86  
 QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQ 92  
 DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQ 117

## RESULT 4

Q61YF0 PRELIMINARY; PRT; 217 AA.

AC Q61YF0  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Growth hormone 1 variant 2.  
 GN Name=GHI;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jorge A.A.L., Arnhold I.J.P., Mendonca B.B.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY613432; AAT11509.1;  
 DR InterPro; IPR009079; 4 helix cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 SQ SEQUENCE 217 AA; 24946 MW; 72D079DF52BD851A CRC64;

Query Match 57.7%; Score 460; DB 2; Length 217;  
 Best Local Similarity 98.9%; Pred. No. 2,1e-36;  
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 61  
 DB 27 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 86  
 QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQ 92  
 DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQ 117

## RESULT 5

AAT11509 PRELIMINARY; PRT; 217 AA.

AC AAT11509  
 DT 20-MAY-2004 (TREMBlrel. 27, Created)  
 DT 20-MAY-2004 (TREMBlrel. 27, Last sequence update)  
 DT 20-MAY-2004 (TREMBlrel. 27, Last annotation update)  
 DE Growth hormone 1 variant 2.  
 GN GHI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Jorge A.A.L., Arnhold I.J.P., Mendonca B.B.;  
 RL "New allelic variant (G152R) in growth hormone (GH) gene associated  
 RT with idiopathic short stature."  
 RI Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY613432; AAT11509.1;  
 SQ SEQUENCE 217 AA; 24946 MW; 72D079DF52BD851A CRC64;

Query Match 57.7%; Score 460; DB 2; Length 217;  
 Best Local Similarity 98.9%; Pred. No. 2,1e-36;  
 Matches 90; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 61  
 DB 27 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 86  
 QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQ 92  
 DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQ 117

## RESULT 6

Q61YF1 PRELIMINARY; PRT; 217 AA.

AC Q61YF1  
 DT 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE Growth hormone 1 variant 1.  
 GN Name=GHI;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jorge A.A.L., Arnhold I.J.P., Mendonca B.B.;  
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY613431; AAT11508.1;  
 DR InterPro; IPR009079; 4 helix cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 SQ SEQUENCE 217 AA; 24875 MW; 12DB1B92F63934D8 CRC64;

Query Match 57.3%; Score 457; DB 2; Length 217;  
 Best Local Similarity 69.7%; Pred. No. 4e-36;  
 Matches 101; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

QY 2 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 61  
 DB 27 FPTPLSLRFDNMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLONPOTSLSFSSES IPT 86  
 QY 62 PSNREETOQKSNLELRISILLIQSWLEPVQ 92  
 DB 87 PSNREETOQKSNLELRISILLIQSWLEPVQ 117

QY 112 ALYIVCG--ERGFYPTKRGIVEQ 134  
 DB 146 GIQLMGRLEDG---SPRTQIFKQ 167

## RESULT 7

AAT11508 PRELIMINARY; PRT; 217 AA.

AC AAT11508  
 DT 20-MAY-2004 (TREMBlrel. 27, Created)  
 DT 20-MAY-2004 (TREMBlrel. 27, Last sequence update)  
 DT 20-MAY-2004 (TREMBlrel. 27, Last annotation update)  
 DE Growth hormone 1 variant 1.  
 GN GHI.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jorge A.A.L., Arnold I.J.P., Mendonça B.B.;  
 RT "New allelic variant (A39V) in growth hormone (GH) gene associated  
 with GH deficiency in heterozygous state."  
 RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF613431; AAT1508.1;  
 SQ SEQUENCE 217 AA; 24875 MW; 12DB1B92F63934D8 CRC64;

Query Match 57.3%; Score 457; DB 2; Length 217;  
 Best Local Similarity 69.7%; Pred. No. 4e-36; Mismatches 21; Indels 16; Gaps 4;  
 Matches 101; Conservative 7;

QY 2 FFTPLSLRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSSESIFT 61  
 DB 27 FFTPLSLRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSSESIFT 86  
 QY 62 PSNREETOQKSNLELRISILLIOSWLEPVQGTGPRFVNOHLCG-----HIVE 111  
 DB 87 PSNREETOQKSNLELRISILLIOSWLEPVQGTGPRFVNOHLCG-----HIVE 145  
 QY 112 ALYIVCG--ERGFYTPKTRGIVEQ 134  
 DB 146 GIQTMGRLEBDG--SPRTGQIFKQ 167

RESULT 8  
 SOMA\_SAIBB STANDARD; PRT; 217 AA.  
 ID SOMA\_SAIBB STANDARD; PRT; 217 AA.  
 AC P58343;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Somatotropin precursor (Growth hormone).  
 GN Name=GH1;  
 OS Saoturi boliviensis boliviensis (Bolivian squirrel monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Saimiri.  
 OX NCBI\_TaxID=39432;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21265430; PubMed=11371582;  
 RT "Epistatic evolution of growth hormone in primates and emergence of the  
 species specificity of human growth hormone receptor."  
 RL Mol. Biol. Evol. 18:945-953(2001).  
 CC -1- FUNCTION: Plays an important role in growth control. Its major  
 role in stimulating body growth is to stimulate the liver and  
 other tissues to secrete IGF-1. It stimulates both the  
 differentiation and proliferation of myoblasts. It also stimulates  
 amino acid uptake and protein synthesis in muscle and other  
 tissues (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the somatotropin/prolactin family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL; AF339060; AAK62287.1; -  
 DR HSR; P01241; 1A22.  
 DR InterPro: IPR009079; 4 helix cytokine.  
 DR InterPro: IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.

DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 KW Hormone; Pituitary; Signal.  
 FT SIGNAL 1 26 By similarity.  
 FT CHAIN 27 217 Somatotropin.  
 FT DISULFD 79 191 By similarity.  
 FT DISULFD 208 215 By similarity.  
 SQ SEQUENCE 217 AA; 24864 MW; 9515289992C529F7 CRC64;

Query Match 54.8%; Score 437; DB 1; Length 217;  
 Best Local Similarity 91.3%; Pred. No. 3.5e-34; Mismatches 3; Indels 0; Gaps 0;  
 Matches 84; Conservative 5;

QY 2 FFTPLSLRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSSESIFT 61  
 DB 27 FFTPLSLRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSSESIFT 86  
 QY 62 PSNREETOQKSNLELRISILLIOSWLEPVQGTGPRFVNOHLCG-----HIVE 111  
 DB 87 PSNREETOQKSNLELRISILLIOSWLEPVQGTGPRFVNOHLCG-----HIVE 145

RESULT 9  
 QSWNEO PRELIMINARY; PRT; 217 AA.  
 ID QSWNEO PRELIMINARY; PRT; 217 AA.  
 AC QSWNEO;  
 DT 01-MAR-2002 (TRENBERL 20, Created)  
 DT 01-MAR-2002 (TRENBERL 20, Last sequence update)  
 DT 01-MAR-2004 (TRENBERL 26, Last annotation update)  
 DE Growth hormone.  
 GN Name=GH-N;  
 OS Ateles geoffroyi (Black-handed spider monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.  
 OX NCBI\_TaxID=9509;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF374234; AAL72286.1; -  
 DR HSR; P01241; 1A22.  
 DR GO; GO:0005176; C:extracellular; IEA.  
 DR GO; GO:0005179; F:hormone activity; IEA.  
 DR InterPro: IPR009079; 4 helix cytokine.  
 DR InterPro: IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 SQ SEQUENCE 217 AA; 24894 MW; 425829F41EBAE6 CRC64;

Query Match 54.7%; Score 436; DB 2; Length 217;  
 Best Local Similarity 66.9%; Pred. No. 4.4e-34; Mismatches 24; Indels 16; Gaps 4;  
 Matches 97; Conservative 8;

QY 2 FFTPLSLRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSSESIFT 61  
 DB 27 FFTPLSLRLFDNAMLRAHRLHQLAFDTYQEFEEAYIPKQKYSFLQNPQTSLSFSSESIFT 86  
 QY 62 PSNREETOQKSNLELRISILLIOSWLEPVQGTGPRFVNOHLCG-----HIVE 111  
 DB 87 PSNREETOQKSNLELRISILLIOSWLEPVQGTGPRFVNOHLCG-----HIVE 145  
 QY 112 ALYIVCG--ERGFYTPKTRGIVEQ 134  
 DB 146 GIQTMGRLEBDG--SPRTGQIFKQ 167

RESULT 10  
 SOM2\_PANTR STANDARD; PRT; 217 AA.  
 ID SOM2\_PANTR STANDARD; PRT; 217 AA.  
 AC P58757;  
 DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, last sequence update)  
 DT 05-JUL-2004 (Rel. 44, last annotation update)  
 DE Growth hormone variant precursor (GH-V) (Placenta-specific growth hormone) (Growth hormone 2).  
 GN Name=GH2;  
 OS Pan troglodytes (Chimpanzee).  
 OC Buiyotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 NCBI\_Taxid=9598;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Revol A., Esquivel D., Santiago D., Barrera-Saldana H.;  
 RT "Independent duplication of the growth hormone gene in three Anthropoid lineages."  
 RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Expressed in the placenta.  
 CC -1- SIMILARITY: Belongs to the somatotropin/prolactin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AF374233; AAL72285.1; -  
 DR HSR; P01241; 1A22.  
 DR InterPro; IPR009079; 4\_helix\_cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 KM Glycoprotein; Hormone; Placenta; Signal.  
 FT SIGNAL 1 26 By similarity.  
 FT CHAIN 27 217 Growth hormone variant.  
 FT DISULFID 79 191 By similarity.  
 FT DISULFID 208 215 By similarity.  
 SQ SEQUENCE 217 AA; 24990 MW; 1592A429075677DE CRC64;  
 Query Match 54.5%; Score 434.5; DB 1; Length 217;  
 Best Local Similarity 78.9%; Pred. No. 6, 1e-34;  
 Matches 90; Conservative 4; Mismatches 9; Indels 11; Gaps 1;  
 QY 2 PPTPLSLRFLPNAMLRARLHQLAFPTYQSEFEAYIPKQKYSFLONPQTSLSFSSSIFT 61  
 DB 27 PPTPLSLRFLPNAMLRARLHQLAFPTYQSEFEAYIPKQKYSFLONPQTSLSFSSSIFT 86  
 QY 62 PSNEETQOKSNLELRISILLIQSWLEPYQL-----GTGPRFVNHQL 104  
 DB 87 PSNVTQOKSNLELRISILLIQSWLEPYQLRSVFANSILVYAGSDSNVVRHL 140  
 RESULT 11  
 ID SOWA\_CALUA STANDARD; PRT; 217 AA.  
 AC Q9GMB3;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Somatotropin precursor (Growth hormone).  
 GN Name=GH1;  
 OS Callithrix jacchus (Common marmoset).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrichx.

CX NCBI\_Taxid=9483;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Wallis O.C., Wallis M.;  
 RT "Cloning and characterization of a putative growth hormone encoding RT gene from the marmoset (Callithrix jacchus)."  
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Plays an important role in growth control. Its major role in stimulating body growth is to stimulate the liver and other tissues to secrete IGF-1. It stimulates both the differentiation and proliferation of myoblasts. It also stimulates amino acid uptake and protein synthesis in muscle and other tissues (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the somatotropin/prolactin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AJ297563; CAC03481.1; -  
 DR HSR; P01241; 1A22.  
 DR InterPro; IPR009079; 4\_helix\_cytokine.  
 DR InterPro; IPR001400; Somatotropin.  
 DR Pfam; PF00103; Hormone\_1; 1.  
 DR PRINTS; PR00836; SOMATOTROPIN.  
 DR PROSITE; PS00266; SOMATOTROPIN\_1; 1.  
 DR PROSITE; PS00338; SOMATOTROPIN\_2; 1.  
 KM Hormone; Pituitary; Signal.  
 FT SIGNAL 1 26 By similarity.  
 FT CHAIN 27 217 Somatotropin.  
 FT DISULFID 79 191 By similarity.  
 FT DISULFID 208 215 By similarity.  
 SQ SEQUENCE 217 AA; 24959 MW; E102151A12CE6192 CRC64;  
 Query Match 54.2%; Score 432; DB 1; Length 217;  
 Best Local Similarity 91.2%; Pred. No. 1, 1e-33;  
 Matches 83; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 PPTPLSLRFLPNAMLRARLHQLAFPTYQSEFEAYIPKQKYSFLONPQTSLSFSSSIFT 61  
 DB 27 PPTPLSLRFLPNAMLRARLHQLAFPTYQSEFEAYIPKQKYSFLONPQTSLSFSSSIFT 86  
 QY 62 PSNEETQOKSNLELRISILLIQSWLEPYQ 92  
 DB 87 PSNVTQOKSNLELRISILLIQSWLEPYQ 117  
 RESULT 12  
 ID Q6FHS4 PRELIMINARY; PRT; 217 AA.  
 AC Q6FHS4;  
 DT 05-JUL-2004 (TRENDArel. 27, Created)  
 DT 05-JUL-2004 (TRENDArel. 27, Last sequence update)  
 DT 05-JUL-2004 (TRENDArel. 27, Last annotation update)  
 DE GH2 protein.  
 GN Name=GH2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_Taxid=9606;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Helleck A., Ebert L., Moundinya M., Schick M., Bisenstein S.,  
 RA Neubert P., Ketting K., Schatten R., Shen B., Henze S., Mar W.,  
 RA Korn B., Zuo D., Hu Y., Labeat J.,  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; CDS41902; CAG46700.1; -  
 DR InterPro; IPR009079; 4\_helix\_cytokine.





```
D8      PSNRVYKTOQKSNIELIRISLIIQSWEIPQLRSVFANSILYYGASDSNVYRHL 140
```

RESULT 15  
ID Q6FH32                  PRELIMINARY; PRT; 217 AA.

```
D9 AC    O6FH32;  
D9 DT   05-JUL-2004 (TRENBLrel. 27, Created)  
D9 DT   05-JUL-2004 (TRENBLrel. 27, Last sequence update)  
D9 DT   05-JUL-2004 (TRENBLrel. 27, Last annotation update)  
DE DE   GH2 protein (Fragment).  
CN CN   Name-GH2;  
OS OS   Homo sapiens (Human).  
OC OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC CC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RX RX   NCBI Taxid=9606;  
RN RN   [1]  
RP RP   SEQUENCE FROM N.A.  
RA RA   Hallack A., Expert L., Mkandinya M., Schick M., Eisenstein S.,  
RA RA   Neubert P., Ketrang X., Schatten R., Shen B., Henze S., Mar W.,  
RL RL   Korn B., Zhuo D., Hu Y., Labaez J.,  
RU RU   Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR DR   EMBL; CRS41924; CAG46722.1; -  
DR DR   InterPro; IPR009079; 4_helix_cytokine.  
DR DR   Pfam; PF00103; Hormone_1; 1.  
DR DR   PRINTS; PR00836; SOMATOTROPIN_1; 1.  
DR DR   PROSITE; PS00266; SOMATOTROPIN_2; 1.  
DR DR   PROSITE; PS00338; SOMATOTROPIN_2; 1.  
FT FT   NON_TER          217  
SQ SQ   SEQUENCE       217 AA; 25010 MW; 075C0E6F3C15AFA5 CRC64;
```

Query Match                      52.9%; Score 421.5; DB 2; Length 217;  
Best Local Similarity        77.2%; Pred.No.1..ie-32;  
Matches     88; Conservative     4; Mismatches    11; Indels    11; Gaps    1;

Cy       2 FETIPLSLFPNMMLEAHRIHQALAFPYOEFEEAYTPKEOKXSFIONPOTSLSFSSESPT 61  
Db       27 FETIPLSRLLFDAMLRARRLYOLAVDTYGFEEAYTLKEOKTSFLONPOTSICPSRESIPT 86

Cy       62 PSNRFEETOQKSNLIELRLISILLIGSWIEPVQT-----GTGPFFVNQH 104  
Db       87 PSNRVYKTOQKSNIELIRISLIIQSWEIPQLRSVFANSILYYGASDSNVYRHL 140

Search completed: November 2, 2004, 20:20:34  
Job time : 145.295 secs